GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 16:07:29 ; Search time 13988 Seconds
(without alignments)
11511.316 Million cell updates/sec

Title: US-09-720-383C-9
Perfect score: 3936
Sequence: 1 cttctccctcgtcggtgcgg.....ctgagttggttttgtgaaaa 3936

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3935	100.0	3968	8	AF200531	AF200531 Zea mays
	2	2543	64.6	3538	8	AF200530	AF200530 Zea mays
	3	2241	56.9	3812	8	AF200532	AF200532 Zea mays
	4	1531.4	38.9	3847	8	AF304374	AF304374 Nicotiana
	5	1509	38.3	3255	6	AX507835	AX507835 Sequence
	6	1489.8	37.9	3444	8	AF062485	AF062485 Arabidops
	7	1434.2	36.4	3771	8	AF027173	AF027173 Arabidops
	8	1431.8	36.4	3255	6	AX505864	AX505864 Sequence
	9	1429.4	36.3	3828	6	AR267558	AR267558 Sequence
	10	1429.4	36.3	3828	6	AX030944	AX030944 Sequence
	11	1429.4	36.3	3828	6	BD022677	BD022677 Manipulat
	12	1378.4	35.0	3725	8	AF200526	AF200526 Zea mays
	13	1364	34.7	3752	8	AF200525	AF200525 Zea mays
c	14	1335.4	33.9	171376	8	AC104487	AC104487 Oryza sat
	15	1332.4	33.9	3763	8	BT008654	BT008654 Arabidops
	16	1330.8	33.8	3673	6	AR267560	AR267560 Sequence
	17	1330.8	33.8	3673	6	AX030948	AX030948 Sequence
	18	1330.8	33.8	3673	6	BD022679	BD022679 Manipulat
	19	1326.6	33.7	3603	6	AR267557	AR267557 Sequence
	20	1326.6	33.7	3603	6	AX030942	AX030942 Sequence
	21	1326.6	33.7	3603	6	BD022676	BD022676 Manipulat
	22	1321.4	33.6	3676	8	AF200529	AF200529 Zea mays
	23	1301.8	33.1	3723	8	AF150630	AF150630 Gossypium
	24	1298.2	33.0	3222	6	AX652952	AX652952 Sequence
c	25	1294.8	32.9	150778	2	AP004298	AP004298 Oryza sat
	26	1289.4	32.8	3745	8	AF200528	AF200528 Zea mays
	27	1279.4	32.5	3795	8	AF200533	AF200533 Zea mays
	28	1279.4	32.5	3799	6	AX338680	AX338680 Sequence
	29	1272.4	32.3	3229	8	BT002335	BT002335 Arabidops
	30	1269.2	32.2	3614	6	AR267559	AR267559 Sequence
	31	1269.2	32.2	3614	6	AX030946	AX030946 Sequence
	32	1269.2	32.2	3614	6	BD022678	BD022678 Manipulat
	33	1269.2	32.2	3682	8	AF027174	AF027174 Arabidops
	34	1268.4	32.2	3264	6	AX653232	AX653232 Sequence
	35	1250.6	31.8	3395	8	AF081534	AF081534 Populus x
	36	1198	30.4	3081	8	AF088917	AF088917 Arabidops
	37	1196.8	30.4	3355	8	AY139754	AY139754 Arabidops
	38	1196.4	30.4	3081	8	BT004543	BT004543 Arabidops
	39	1192.6	30.3	3277	8	AY095297	AY095297 Populus t
	40	1192.4	30.3	3181	8	BT006111	BT006111 Arabidops
	41	1192.4	30.3	3377	8	BT005710	BT005710 Arabidops
	42	1184	30.1	3168	8	AF458083	AF458083 Arabidops
	43	1171.2	29.8	2828	8	AF200527	AF200527 Zea mays
	44	1170	29.7	2830	6	AX338676	AX338676 Sequence
c	45	1163.4	29.6	130300	8	AP005824	AP005824 Oryza sat

ALIGNMENTS

RESULT 1
AF200531
LOCUS AF200531 3968 bp mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-7 (CesA-7) mRNA, complete cds.
ACCESSION AF200531
VERSION AF200531.1 GI:9622885
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3968)
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
Xoconostle-Cazares,B. and Delmer,D.P.

```
  TITLE     A comparative analysis of the plant cellulose synthase (CesA) gene
            family
  JOURNAL   Plant Physiol. 123 (4), 1313-1324 (2000)
  MEDLINE   20398328
   PUBMED   10938350
REFERENCE   2  (bases 1 to 3968)
  AUTHORS   Dhugga,K.S. and Helentjaris,T.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
            International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES             Location/Qualifiers
     source          1. .3968
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="taxon:4577"
                     /note="identified and isolated from the maize genome
                     database cDNA collection of Pioneer Hi-Bred International,
                     Inc"
     gene            1. .3968
                     /gene="CesA-7"
     CDS             143. .3403
                     /gene="CesA-7"
                     /note="CESA-7"
                     /codon_start=1
                     /product="cellulose synthase-7"
                     /protein_id="AAF89967.1"
                     /db_xref="GI:9622886"
                     /translation="MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGD
                     DVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEE
                     EDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQ
                     MVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWK
                     ERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMII
                     IIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLD
                     RLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSD
                     DGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRER
                     RAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGG
                     LDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINN
                     SKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQG
                     PIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTE
                     KKKLLFPKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLL
                     ENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGW
                     RSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLER
                     FSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSIL
                     EMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEF
                     SELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHL
                     YPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN"
BASE COUNT     1022 a    837 c   1018 g   1091 t
ORIGIN

  Query Match           100.0%;  Score 3935;  DB 8;  Length 3968;
  Best Local Similarity 100.0%;  Pred. No. 0;
  Matches 3935;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGGA 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGGA 60

Qy         62 TGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCGG 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCGG 120

Qy        122 CGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAA 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAA 180

Qy        182 CCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGA 240

Qy        242 GCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGA 300

Qy        302 CCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGA 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGA 360

Qy        362 GCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGG 420

Qy        422 CTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTT 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTT 480

Qy        482 CAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAG 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAG 540

Qy        542 CTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGT 600

Qy        602 TCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGT 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGT 660

Qy        662 GCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAG 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAG 720

Qy        722 CTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGG 780

Qy        782 TAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCA 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCA 840

Qy        842 GACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGA 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGA 900

Qy        902 AGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAG 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAG 960

Qy        962 GATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCA 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCA 1020

Qy       1022 TCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGC 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGC 1080

Qy       1082 CATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCT 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCT 1140

Qy       1142 AGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGA 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGA 1200

Qy       1202 TTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGT 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGT 1260

Qy       1262 TCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGA 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGA 1320

Qy       1322 TGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATG 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATG 1380

Qy       1382 GGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACA 1441
```

```
Db      1381 GGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACA 1440
Qy      1442 GAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAAT 1501
Db      1441 GAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAAT 1500
Qy      1502 GAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAA 1561
Db      1501 GAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAA 1560
Qy      1562 AGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCG 1621
Db      1561 AGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCG 1620
Qy      1622 TGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGG 1681
Db      1621 TGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGG 1680
Qy      1682 AAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCA 1741
Db      1681 AAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCA 1740
Qy      1742 TAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCC 1801
Db      1741 TAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCC 1800
Qy      1802 ATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGC 1861
Db      1801 ATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGC 1860
Qy      1862 AATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCA 1921
Db      1861 AATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCA 1920
Qy      1922 AAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGA 1981
Db      1921 AAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGA 1980
Qy      1982 TATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGT 2041
Db      1981 TATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGT 2040
Qy      2042 ATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAG 2101
Db      2041 ATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAG 2100
Qy      2102 GACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCA 2161
Db      2101 GACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCA 2160
Qy      2162 AAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGA 2221
Db      2161 AAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGA 2220
Qy      2222 GAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAA 2281
Db      2221 GAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAA 2280
Qy      2282 TGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGT 2341
Db      2281 TGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGT 2340
Qy      2342 TTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTC 2401
Db      2341 TTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTC 2400
Qy      2402 TCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGG 2461
Db      2401 TCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGG 2460
Qy      2462 AAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGAT 2521
Db      2461 AAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGAT 2520
Qy      2522 GCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTC 2581
Db      2521 GCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTC 2580
Qy      2582 TGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTAT 2641
Db      2581 TGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTAT 2640
Qy      2642 TGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATT 2701
Db      2641 TGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATT 2700
Qy      2702 TTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTT 2761
Db      2701 TTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTT 2760
Qy      2762 GGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCT 2821
Db      2761 GGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCT 2820
Qy      2822 GAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCAT 2881
Db      2821 GAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCAT 2880
Qy      2882 CCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTG 2941
Db      2881 CCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTG 2940
Qy      2942 GGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCAT 3001
Db      2941 GGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCAT 3000
Qy      3002 AGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTC 3061
Db      3001 AGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTC 3060
Qy      3062 AGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATT 3121
Db      3061 AGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATT 3120
Qy      3122 GAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATG 3181
Db      3121 GAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATG 3180
Qy      3182 GGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTT 3241
Db      3181 GGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTT 3240
Qy      3242 TCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCAT 3301
Db      3241 TCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCAT 3300
Qy      3302 TCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGA 3361
Db      3301 TCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGA 3360
Qy      3362 TGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGC 3421
Db      3361 TGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGC 3420
Qy      3422 TCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGTC 3481
Db      3421 TCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGTC 3480
Qy      3482 TGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAAT 3541
Db      3481 TGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAAT 3540
Qy      3542 ATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCGT 3601
Db      3541 ATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCGT 3600
```

```
Qy       3602 GGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATAT 3661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATAT 3660

Qy       3662 AGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACTT 3721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACTT 3720

Qy       3722 GGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATTA 3781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATTA 3780

Qy       3782 TTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAAA 3841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAAA 3840

Qy       3842 ACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGTT 3901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 ACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGTT 3900

Qy       3902 GTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             ||||||||||||||||||||||||||||||||||||
Db       3901 GTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3935
```

```
RESULT 2
AF200530
LOCUS       AF200530               3538 bp    mRNA    linear   PLN 31-AUG-2000
DEFINITION  Zea mays cellulose synthase-6 (CesA-6) mRNA, complete cds.
ACCESSION   AF200530
VERSION     AF200530.1  GI:9622883
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3538)
  AUTHORS   Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
            Xoconostle-Cazares,B. and Delmer,D.P.
  TITLE     A comparative analysis of the plant cellulose synthase (CesA) gene
            family
  JOURNAL   Plant Physiol. 123 (4), 1313-1324 (2000)
  MEDLINE   20398328
   PUBMED   10938350
REFERENCE   2  (bases 1 to 3538)
  AUTHORS   Dhugga,K.S. and Helentjaris,T.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
            International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES             Location/Qualifiers
     source          1. .3538
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="taxon:4577"
                     /note="identified and isolated from the maize genome
                     database cDNA collection of Pioneer Hi-Bred International,
                     Inc"
     gene            1. .3538
                     /gene="CesA-6"
     CDS             43. .3222
                     /gene="CesA-6"
                     /note="CESA-6"
                     /codon_start=1
                     /product="cellulose synthase-6"
                     /protein_id="AAF89966.1"
                     /db_xref="GI:9622884"
                     /translation="MDQRNGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYEYER
                     REGTQNCPQCKTRFKRFKGCARVPGDEEEDGVDDLENEFNWSDKHDSQYLAESMLHAH
                     MSYGRGADLDGVPQPFHPIPNVPLLTNGQMVDDIPPDQHALVPSFVGGGGKRIHPLPY
                     ADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMESWKQKQERMHQTRNDGGGDDGDDAD
                     LPLMDEARQPLSRKIPLPSSQINPYRMIIIIRLVVLCFFFHYRVMHPVPDAFALWLIS
                     VICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGHPSQLAPVDFFVSTVDPLK
                     EPPLVTANTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYS
                     LEPRAPEWYFQQKIDYLKDKVAPNFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW
                     TMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDVEGNELPRLVYVSREKRPGYNHHKKAG
                     AMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRF
                     DGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSR
                     TCNCWPKWCICCCCFGNRKTKKKTKTSKPKFEKIKKLFKKKENQAPAYALGEIDEAAP
                     GAENEKASIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYE
                     DKTGWGKDIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRAAFKGSAPLNLSDRFHQV
                     LRWALGSIEILFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICL
                     LTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSH
                     LFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVV
                     AGISNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLAS
                     IFSLLWVRIDPFLAKDDGPLLEECGLDCN"
BASE COUNT      876 a    813 c    928 g    920 t      1 others
ORIGIN
```

Query Match 64.6%; Score 2543; DB 8; Length 3538;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 415; Indels 9; Gaps 2;

```
Qy        192 CTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGG 251
              ||||||||||||||||||| ||||| || |||||||||||    |    |||| |||||||
Db          1 CTCGTCGTCATCCGCCGCGATGGCGAGCCAGGGCCGAAGCCCATGGACCAGCGGAACGGC 60

Qy        252 CAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTG 311
              |||||||||||||||||||||||||||| || |    |||||  |||||| || ||||||
Db         61 CAGGTGTGCCAGATTTGCGGCGACGACGTGGGGCGCAACCCCGACGGGGAGCCTTTCGTG 120

Qy        312 GCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAG 371
              || ||||||||||||||||||||||| |||||||||||||||||||| |||||||||||| |||
Db        121 GCCTGCAACGAGTGCGCCTTCCCCATCTGCCGGGACTGCTACGAGTACGAGCGCCGCGAG 180

Qy        372 GGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGT 431
              |||||||||||||||||||||||||||| || | ||||||| ||||||| |||   ||
Db        181 GGCACGCAGAACTGCCCCCAGTGCAAGACCCGCTTCAAGCGCTTCAAGGGGTGCGCGCGC 240

Qy        432 GTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGG--- 488
              ||| |||| |||||||||||||||||||||| |||||||| ||||||||||||||
Db        241 GTGCCCGGGGACGAGGAGGAGGACGGCGTCGACGACCTGGAGAACGAGTTCAACTGGAGC 300

Qy        489 GACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGC 548
              |||   || ||||| ||||   | |||||||||||||| ||| ||||||||||||||||
Db        301 GACAAGCACGACTCCCAGTACCTCGCCGAGTCCATGCTCCACGCCCACATGAGCTACGGC 360

Qy        549 CGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTC 608
              || || |  ||||   | |||| ||| ||  | ||||| | || ||||||||||||| |||
Db        361 CGCGGCGCCGACCTCGACGGCGTGCCGCAGCCATTCCACCCCATCCCCAATGTTCCCCTC 420

Qy        609 CTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCT 668
              |||||||||| || ||||| |||||||||| ||||| |||||||| || ||||| ||
Db        421 CTCACCAACGGACAGATGGTCGATGACATCCCGCCGGACCAGCACGCCCTTGTGCCCTCG 480

Qy        669 TTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCT 728
              ||| ||||||| |||||  |||||||| || || || || || |||||||||| | | |||
Db        481 TTCGTGGGTGGCGGGGGGAAGAGGATTCACCCTCTCCCGTACGCGGATCCCAACCTTCCT 540

Qy        729 GTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTT 788
              |||||||| |||||||||||||| |||||||||| || |||||||| || || || ||
Db        541 GTGCAACCGAGGTCTATGGACCCTTCCAAGGATCTCGCCGCATATGGCTACGGGAGCGTA 600

Qy        789 GCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGG 848
              ||.|||||||  |||||||||  |||||||||  || |||||||||||||||||||| ||
Db        601 GCATGGAAGGAGAGGATGGAGAGCTGGAAGCAGAAGCAGGAGAGGATGCACCAGACGAGG 660

Qy        849 AATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGA 908
              || ||||| || || |||||||||| ||||| |||||||||||||||||||||| |||
Db        661 AACGATGGCGGCGGCGATGATGGTGATGATGCAGATCTACCACTAATGGATGAAGCTAGA 720

Qy        909 CAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATT 968
              || | | ||||||| || || || ||||| |||||||| || || || ||||||||||||
Db        721 CAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATCAACCCCTATAGGATGATT 780
```

```
Qy       969 ATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTG 1028
Db       781 ATAATAATTCGGCTAGTGGTTTTGTGTTTCTTCTTCCACTACCGAGTGATGCATCCGGTG 840

Qy      1029 AATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCT 1088
Db       841 CCTGATGCATTTGCTTTATGGCTCATATCTGTGATCTGTGAAATTTGGTTTGCCATGTCT 900

Qy      1089 TGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGG 1148
Db       901 TGGATTCTTGACCAGTTTCCAAAGTGGTTTCCTATCGAGAGGGAAACCTATCTTGACCGG 960

Qy      1149 CTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTT 1208
Db       961 CTGAGTTTAAGGTTTGACAAGGAAGGGCATCCTTCTCAACTCGCCCCTGTTGATTTCTTT 1020

Qy      1209 GTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCT 1268
Db      1021 GTCAGTACGGTTGATCCCTTGAAGGAACCTCCATTGGTCACTGCTAATACTGTTCTATCT 1080

Qy      1269 ATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCT 1328
Db      1081 ATCCTTTCGGTGGATTATCCAGTTGATAAGGTTTCATGCTACGTTTCTGATGATGGTGCT 1140

Qy      1329 GCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCT 1388
Db      1141 GCCATGCTGACATTTGAAGCATTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCT 1200

Qy      1389 TTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATA 1448
Db      1201 TTCTGCAAAAGATATAGCCTTGAGCCTCGTGCTCCAGAGTGGTACTTCCAACAGAAGATA 1260

Qy      1449 GACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGA 1508
Db      1261 GACTACCTGAAAGACAAGGTGGCGCCAAACTTTGTTAGAGAACGGAGAGCAATGAAGAGA 1320

Qy      1509 GAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCT 1568
Db      1321 GAGTATGAGGAATTCAAGGTCAGAATCAATGCCTTGGTTGCTAAAGCCCAAAAGGTTCCT 1380

Qy      1569 GAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCAT 1628
Db      1381 GAGGAAGGATGGACAATGCAGGATGGAACTCCATGGCCCGGAAATAATGTCCGTGATCAT 1440

Qy      1629 CCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAA 1688
Db      1441 CCTGGAATGATTCAGGTTTTCCTTGGTCAAAGTGGTGGCCATGATGTGGAAGGAAATGAG 1500

Qy      1689 CTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAA 1748
Db      1501 CTGCCTCGATTGGTTTATGTTTCAAGAGAAAAACGGCCAGGCTACAACCATCACAAGAAG 1560

Qy      1749 GCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTG 1808
Db      1561 GCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACTAATGCTCCTTATTTG 1620

Qy      1809 TTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGT 1868
Db      1621 CTGAACTTGGATTGTGATCACTATATCAATAATAGTAAGGCTATAAAGGAAGCAATGTGT 1680

Qy      1869 TTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTT 1928
Db      1681 TTTATGATGGATCCTTTGCTTGGAAAGAAAGTTTGCTATGTGCAGTTTCCTCAAAGATTT 1740

Qy      1929 GATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTTGATATCAAC 1988
Db      1741 GATGGGATTGATCGCCATGATCGATATGCTAACAGAAATGTTGTCTTTTTCGATATCAAC 1800

Qy      1989 ATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGA 2048
Db      1801 ATGAAAGGTTTGGATGGTATCCAGGGCCCAATTTATGTGGGTACTGGATGTGTCTTCAGA 1860

Qy      2049 AGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGC 2108

Db      1861 AGGCAGGCATTATATGGCTACGATGCTCCCAAAACAAAGAAGCCACCATCAAGAACTTGC 1920

Qy      2109 AACTGCTGGCCCAAGTGGTGCTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAG 2168
Db      1921 AACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTTGGTAACAGGAAGACCAAGAAG 1980

Qy      2169 ACTACCAAACCCAAAACAGAGA------AGAAAAAGTTATTATTTTTCAAGAAAGAAGAG 2222
Db      1981 AAGACCAAGACCTCTAAACCTAAATTTGAGAAGATAAAGAAACTTTTTAAGAAAAAGGAA 2040

Qy      2223 AACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAAT 2282
Db      2041 AATCAAGCCCCTGCATATGCTCTTGGTGAAATTGATGAAGCCGCTCCAGGAGCTGAAAAT 2100

Qy      2283 GAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTT 2342
Db      2101 GAAAAGGCTAGTATTGTAAATCAACAGAAGTTGGAAAAGAAATTTGGCCAGTCTTCAGTT 2160

Qy      2343 TTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCT 2402
Db      2161 TTTGTTGCATCCACACTTCTTGAGAATGGTGGAACCCTGAAGAGTGCCAGTCCAGCTTCT 2220

Qy      2403 CTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGA 2462
Db      2221 CTTCTGAAGGAAGCTATACATGTCATCAGTTGTGGATATGAAGACAAAACAGGCTGGGGA 2280

Qy      2463 AAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATG 2522
Db      2281 AAAGATATTGGTTGGATTTATGGATCAGTCACAGAAGATATTCTTACTGGGTTTAAGATG 2340

Qy      2523 CATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCT 2582
Db      2341 CACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGCCGCCTTCAAAGGTTCC 2400

Qy      2583 GCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATT 2642
Db      2401 GCACCTCTCAATCTTTCCGATCGTTTTCACCAGGTTCTTCGGTGGGCTCTTGGTTCAATT 2460

Qy      2643 GAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTT 2702
Db      2461 GAAATTTTGTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGTGGTGGACTAAAGTTC 2520

Qy      2703 TTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTG 2762
Db      2521 CTGGAAAGGTTTTCGTACATTAACTCCATCGTATACCCTTGGACATCTATCCCGCTCTTG 2580

Qy      2763 GCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTG 2822
Db      2581 GCCTATTGCACATTGCCTGCCATCTGCTTGCTGACAGGGAAATTTATCACGCCAGAGCTT 2640

Qy      2823 AATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATC 2882
Db      2641 AACAATGTTGCCAGCCTCTGGTTCATGTCACTTTTCATCTGCATTTTTGCTACGAGCATC 2700

Qy      2883 CTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGG 2942
Db      2701 CTGGAAATGAGATGGAGTGGTGTAGGCATCGATGACTGGTGGAGAAACGAGCAGTTTTGG 2760

Qy      2943 GTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATA 3002
Db      2761 GTCATTGGAGGCGTGTCTTCACATCTCTTTGCTGTGTTCCAGGGACTCCTCAAGGTCATA 2820

Qy      3003 GCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCA 3062
Db      2821 GCTGGTGTAGACACGAGCTTCACTGTGACATCCAAGGGCGGAGACGACGAGGAGTTCTCA 2880

Qy      3063 GAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTG 3122
Db      2881 GAGCTGTACACATTCAAATGGACGACCCTTCTGATACCTCCGACAACCCTGCTCCTACTG 2940

Qy      3123 AACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGG 3182
```

```
Db       2941 AACTTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAACGGATATGAATCATGG 3000

Qy       3183 GGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTT 3242
             |||||||| || |||||||| |||||||||||||||||||| |||||||||| || ||
Db       3001 GGCCCCCTGTTCGGGAAGCTCTTCTTTGCATTTTGGGTGATCGTCCATCTTTACCCGTTC 3060

Qy       3243 CTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATT 3302
             ||||| ||| ||||||| ||||| |||||||| |||||||||||||| ||||||||||
Db       3061 CTCAAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTGTCATTGTCTGGTCCATC 3120

Qy       3303 CTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGAT 3362
             || |||||||| |||||||||| |||||||| ||||| || || ||||||||||||||||
Db       3121 CTCCTGGCTTCGATCTTCTCGCTGCTTTGGGTCCGGATCGACCCGTTCCTTGCGAAGGAT 3180

Qy       3363 GATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGCT 3422
             |||||||| ||   | |||||||||||| |||||||||||||||| |||||  | |   ||
Db       3181 GATGGTCCCCTGTTGGAGGAGTGTGGTCTGGATTGCAACTAGGAGGTCAGCACGTGGACT 3240

Qy       3423 CCCCC 3427
              ||||
Db       3241 TCCCC 3245
```

```
RESULT 3
AF200532
LOCUS       AF200532               3812 bp    mRNA    linear   PLN 31-AUG-2000
DEFINITION  Zea mays cellulose synthase-8 (CesA-8) mRNA, complete cds.
ACCESSION   AF200532
VERSION     AF200532.1  GI:9622887
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3812)
  AUTHORS   Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
            Xoconostle-Cazares,B. and Delmer,D.P.
  TITLE     A comparative analysis of the plant cellulose synthase (CesA) gene
            family
  JOURNAL   Plant Physiol. 123 (4), 1313-1324 (2000)
  MEDLINE   20398328
   PUBMED   10938350
REFERENCE   2  (bases 1 to 3812)
  AUTHORS   Dhugga,K.S. and Helentjaris,T.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
            International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES             Location/Qualifiers
     source          1. .3812
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="taxon:4577"
                     /note="identified and isolated from the maize genome
                     database cDNA collection of Pioneer Hi-Bred International,
                     Inc"
     gene            1. .3812
                     /gene="CesA-8"
     CDS             214. .3498
                     /gene="CesA-8"
                     /note="CESA-8"
                     /codon_start=1
                     /product="cellulose synthase-8"
                     /protein_id="AAF89968.1"
                     /db_xref="GI:9622888"
                     /translation="MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAPCQIC
                     GDEVGVGFDGEPFVACNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGD
                     EEEDGVDDLEGEFGLQDGAAHEDDPQYVAESMLRAQMSYGRGGDAHPGFSPVPNVPLL
                     TNGQMVDDIPPEQHALVPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYG
                     YGSVAWKERMEGWKQKQERLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSR
                     INPYRMIIVIRLVVLGFFFHYRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLP
                     IERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVE
                     KVSCYVSDDGAAMLTFEALSETSEFAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKV
                     AASFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQ
                     VFLGQSGGRDVEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLSNAAYLLNL
                     DCDHYINNSKAIKEAMCFMMDPLVGKKVCYVQFPQRFDGIDKNDRYANRNVVFFDINM
                     KGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCLSCCCSRNKNKK
                     KTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGAPGADIEKAGIVNQQKLEKKFGQSS
                     VFVASTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSITEDILTG
                     FKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRWALGSVEIFFSKHCPLWYGYG
                     GGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELTNVASIWFMALFI
                     CISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQGLLKVFAGIDTSFTVTS
                     KAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINNGYESWGPLFGKLFF
                     AFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDPFLAKSNGPLLEE
                     CGLDCN"
BASE COUNT     873 a    966 c   1111 g    862 t
ORIGIN
```

Query Match 56.9%; Score 2241; DB 8; Length 3812;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| |||||||||||||||| || |||
Db        193 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 252

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || |||||||| || || |||||||||||| |||| ||   | ||||   | || ||  |
Db        253 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 312

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||    |||||      | ||||||||| |||||||||| ||||||  | | |   || |
Db        313 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 372

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        373 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 432

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||| |||||| ||||    ||||| ||||||| ||| || ||||||||||||
Db        433 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 492

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              |||||||||| || ||| |||| |||||||||||||||||||||||||| ||||||||  |
Db        493 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 552

Qy        477 GAGTTCAACTGG----------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              ||||||  |  |                |||   | ||| ||||||  || |||||||||
Db        553 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 612

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
              ||||||   |  || |||||||||||||| || || ||          |||||  |  |
Db        613 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 663

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
              |||   | |  |||||| || || |||||||||||||| || ||||| ||||||||||
Db        664 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 723

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
              |||||||||||||||| ||||| || | ||||         || || ||||| ||||||
Db        724 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 783

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
              || || || || ||||  || ||||||| | | || |||||||| || || ||||||||
Db        784 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 843

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
              |||||||||| || || || || ||||| || || || ||||||||  | ||||||
Db        844 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 903

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
              |||||||||| || ||| |  |||| ||     ||| || ||||| ||||||   |||
Db        904 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 963

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
              || |||||||| ||||| |||||||||||||||| || || | | ||||||| ||| |
Db        964 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 1023
```

```
Db        964 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1023

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989

Db       1024 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1083

Qy        990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049

Db       1084 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1143

Qy       1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109

Db       1144 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1203

Qy       1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169

Db       1204 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1263

Qy       1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229

Db       1264 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1323

Qy       1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289

Db       1324 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1383

Qy       1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349

Db       1384 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1443

Qy       1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409

Db       1444 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1503

Qy       1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469

Db       1504 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1563

Qy       1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529

Db       1564 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1623

Qy       1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589

Db       1624 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1683

Qy       1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649

Db       1684 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1743

Qy       1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709

Db       1744 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1803

Qy       1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769

Db       1804 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1863

Qy       1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829

Db       1864 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1923

Qy       1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889

Db       1924 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1983

Qy       1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949

Db       1984 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2043

Qy       1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009

Db       2044 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2103
```

```
Qy       2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069

Db       2104 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2163

Qy       2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129

Db       2164 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2223

Qy       2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189

Db       2224 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2283

Qy       2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249

Db       2284 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2343

Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309

Db       2344 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2403

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369

Db       2404 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2463

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429

Db       2464 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2523

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489

Db       2524 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2583

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549

Db       2584 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2643

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609

Db       2644 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2703

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669

Db       2704 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2763

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729

Db       2764 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2823

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789

Db       2824 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2883

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849

Db       2884 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2943

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909

Db       2944 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3003

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969

Db       3004 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3063

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029

Db       3064 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3123

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089

Db       3124 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3183
```

```
Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
              |  |||||| || |||||  | ||  ||  |||||||||| || ||||| || ||  |
Db       3184 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3243

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
              ||  || |||||||| ||||| || ||||| |||||||||||| || |||||||| |||||
Db       3244 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3303

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
              ||  || |||||||| ||||| ||  ||  ||  || ||||| ||| |||| ||  ||||| |||
Db       3304 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3363

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
              |||||  ||  ||||| |||||||||||||||| |||||||| ||  ||||||||||||||
Db       3364 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3423

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
              |||||  ||  | ||  ||  ||||| ||  |||     | ||  ||||| ||  ||||||||||
Db       3424 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3483

Qy       3390 TTGGATTGCAACT 3402
              ||||  |||||||
Db       3484 CTGGACTGCAACT 3496
```

```
RESULT 4
AF304374
LOCUS       AF304374               3847 bp    mRNA    linear   PLN 02-MAY-2001
DEFINITION  Nicotiana alata cellulose synthase catalytic subunit (CesA1) mRNA,
            complete cds.
ACCESSION   AF304374
VERSION     AF304374.1  GI:13925880
KEYWORDS    .
SOURCE      Nicotiana alata (Persian tobacco)
  ORGANISM  Nicotiana alata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 3847)
  AUTHORS   Doblin,M.S., De Melis,L., Newbigin,E., Bacic,A. and Read,S.M.
  TITLE     Pollen tubes of Nicotiana alata express two genes from different
            beta-glucan synthase families
  JOURNAL   Plant Physiol. 125 (4), 2040-2052 (2001)
  MEDLINE   21196092
   PUBMED   11299383
REFERENCE   2  (bases 1 to 3847)
  AUTHORS   Doblin,M.S., De Melis,L., Newbigin,E., Bacic,A. and Read,S.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Plant Biology, University of California, 1
            Shields Ave, Davis, CA 95616, USA
FEATURES             Location/Qualifiers
     source          1. .3847
                     /organism="Nicotiana alata"
                     /mol_type="mRNA"
                     /db_xref="taxon:4087"
                     /tissue_type="pistil"
                     /note="genotype: S2S2"
     gene            1. .3847
                     /gene="CesA1"
     CDS             103. .3378
                     /gene="CesA1"
                     /codon_start=1
                     /product="cellulose synthase catalytic subunit"
                     /protein_id="AAK49454.1"
                     /db_xref="GI:13925881"
                     /translation="MDTKGRLVAGSHNRNEFVVINADDVGRVTSVKELSGQICQICGD
                     EIEVTVDGEPFIACNECAFPVCRQCYEYERREGNQACPQCKTRFKRIKGSPRVDGDDE
                     DDEFDDLDHEFDYHGNPRYMSEAAFSSRLGRGTNHNASGLTTPSEVDPAALNSEIPLL
                     TYGQEDDTISADKHALIIPPFMGRGKKVHPVPYSDSMSLPPRPMDPKKDLAVYGYGTV
                     AWKERMEDWKKKQNDKLQVVKHGGGKGGGNDGDELDDPDLPKMDEGRQPLSRKLPISS
                     SRLSPYRLLILVRLAVVGLFFHYRITHPVNDAYALWLISIICEIWFAVSWIFDQFPKW
                     FPIVRETYLDRLSLRYEKEGKPSGLAPIDIFVSTVDPLKEPPLITANTVLSILAVDYP
                     EDKVSCYVSDDGAAMLTFEALSETSEFARKWVPFCKKFNIEPRAPEWYFSQKVDYLKN
                     KVHPSFVRERRAMKRDYEEFKVRINGLVATAQKVPEDGWTMQDGTPWPGNLVRDHPGM
                     IQVFLGNDGVRDIEGNVLPRLIYVSREKRPGFDHHKKAGAMNALMRVSAVISNAPYLL
                     NVDCDHYINNSKALREAMCFMMDPTSGKKICYVQFPQRFDGIDRHDRYSNRNVVFFDI
                     NMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKKTKPPGKTCNCWPKWCCCCFGSRKKH
                     KKGKTTKDNKKKTKTKEASPQIHALENIEEGIEGIDSEKATLMPQIKLEKKFGQSPVF
                     VASTLLEDGGIPPGATSASLLKEAIHVISCGYEDKTEWGREVGWIYGSVTEDILTGFK
                     MHCHGWRSVYCMPKRPAFKGSAPINLSDRLHQVLRWALGSVEILLSKHCPIWYGYGCG
                     LKPLERFSYINSVVYPLTSLPLIAYCALPAVCLLTGKFIVPEISNYASILFMGLFIMI
                     AATSVLEMQWGGVTIDDWWRNEQFWVIGGASSHLFALFQGLLKVLAGVSTSFTVTSKA
                     ADDGEFSELYLFKWTSLLIPPMTLLIINIIGVIVGISDAINNGYDSWGPLFGRLFFAL
                     WVIVHLYPFLKGVMGRQNKVPTIIVVWSILLASIFSLLWVRVNPFTARGGLVLEVCGL
                     DCE"
BASE COUNT     1061 a    719 c    925 g   1142 t
ORIGIN
```

Query Match 38.9%; Score 1531.4; DB 8; Length 3847;
Best Local Similarity 68.2%; Pred. No. 8.8e-286;
Matches 2252; Conservative 0; Mismatches 996; Indels 53; Gaps 7;

```
Qy        139 AAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCG 198
              |||  |||||  | |  |  ||||  || || || || |||||  | || |||  | || |
Db         98 AAGTAATGGATACTAAAGGGAGGCTAGTTGCTGGTTCACACAATAGGAATGAGTTTGTTG 157

Qy        199 TCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGT 258
              |||||       || |  | |   |   |   |   ||   |  |||||| | |
Db        158 TCATCAATGCTGATGATGTTGGAAGAGTAACTTCTGTGAAAGAATTAAGTGGGCAGATTT 217

Qy        259 GCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCA 318
              |||||||||| || || ||  | |   || |    |  ||||| || ||  | || ||||
Db        218 GCCAGATTTGTGGAGATGAGATTGAAGTTACAGTAGATGGGGAGCCATTTATTGCTTGCA 277

Qy        319 ACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGC 378
              ||||  ||  || |||||  || ||| |  | |||||  ||  ||  ||   |  | ||  ||  |  |
Db        278 ACGAATGTGCATTCCCTGTTTGCAGACAATGCTATGAGTATGAAAGGAGAGAAGGAAATC 337

Qy        379 AGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCG 438
              |    ||||| || ||||| ||||| | |||||||  | ||  ||  |||   | ||    |
Db        338 AAGCTTGCCCTCAATGCAAAACTCGGTTCAAGCGCATTAAAGGGAGCCCCAGGGTTGATG 397

Qy        439 GTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATG 498
              | || || || || || |   | |||||| | ||| | ||||||  | |   | ||  |
Db        398 GAGATGATGAAGATGATGAATTTGATGACTTAGACCATGAGTTCGATTACCATGG-TAAC 456

Qy        499 ACTCGCAGTCTGTGGCCGAGTCCATGCTCTAC-------GGCCACATGAGCTACGGCCGT 551
              ||||   | |||    |  | ||| |||| |       ||||  |  | | |       |
Db        457 CCTCGATATATGTCTGAGGCTGCATTCTCTTCTCGTCTTGGCCGCGGTACCAATCATAAT 516

Qy        552 GGAGGTGACCCTAATGGCGCGCCACAAG------CTTTCCAGCTCAACCCCAATGTTCCA 605
              |    ||  | || |  | |  || |||      ||      || ||| |  |  | ||
Db        517 GCTTCTGGGCTTACTACCCCATCAGAAGTAGATCCTGCTGCTCTTAACTCAGAGATCCCT 576

Qy        606 CTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCT 665
              || || ||  | || |||  ||  |||   ||  |  |  | |||| || ||  |   |
Db        577 CTTCTTACTTATGGTCAAGAGGATGATACAATTTCCGCTGACAAGCATGCTCTTATCATT 636

Qy        666 TCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTA 725
              |    |    ||| | |||||||   | |||||  | ||||||  |||  |||    |
Db        637 CCACCGTTTATGGGTCGCGGAAAGAAAGTTCATCCTGTGCCTTATTCTGATTCCA---TG 693

Qy        726 CCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGT 785
              || |||  ||  |  | ||||| ||    || ||| | || |  |||| |||| | |
Db        694 TCTTTGCCCCCTCGACCCATGGATCCTAAGAAAGATTTGGCAGTTTATGGATATGGGACT 753

Qy        786 GTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACG 845
              ||||| |||||||| | |||||| | |||||| | | |||  |   |     |||  |
Db        754 GTTGCATGGAAGGAAAGAATGGAGGACTGGAAGAAAAAACAGAATGATAAATTACAGGTG 813

Qy        846 GGGAATGATGGTGGTGGTGATGATGGT--------------------GACGATGCTGAT 884
              |  ||  | || || ||| | | ||||                    || ||| | |||
Db        814 GTTAAGCACGGAGGCGGTAAAGGTGGTGGTAACGATGGAGATGAGCTGGATGATCCCGAT 873

Qy        885 CTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGC 944
              | ||    |||||||||||  |||||||  ||| || || ||| | ||  || | || |||
```

```
Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
              |  ||||||| || |||||  | || ||  |||||||||| || ||||| || ||  |
Db       3184 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3243

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
              || || |||||||| ||||| || ||||| |||||||||||| || |||||||| |||||
Db       3244 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3303

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
              || || |||||||| ||||| || || || || ||||| ||| |||| || ||||| |||
Db       3304 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3363

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
              ||||| || ||||| ||||||||||||||||| |||||||| || ||||||||||||||
Db       3364 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3423

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
              |||||  ||  | || || ||||| || |||    | || ||||| || ||||||||||
Db       3424 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3483

Qy       3390 TTGGATTGCAACT 3402
              |||| |||||||
Db       3484 CTGGACTGCAACT 3496
```

```
RESULT 4
AF304374
LOCUS       AF304374               3847 bp    mRNA    linear   PLN 02-MAY-2001
DEFINITION  Nicotiana alata cellulose synthase catalytic subunit (CesA1) mRNA,
            complete cds.
ACCESSION   AF304374
VERSION     AF304374.1  GI:13925880
KEYWORDS    .
SOURCE      Nicotiana alata (Persian tobacco)
  ORGANISM  Nicotiana alata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 3847)
  AUTHORS   Doblin,M.S., De Melis,L., Newbigin,E., Bacic,A. and Read,S.M.
  TITLE     Pollen tubes of Nicotiana alata express two genes from different
            beta-glucan synthase families
  JOURNAL   Plant Physiol. 125 (4), 2040-2052 (2001)
  MEDLINE   21196092
   PUBMED   11299383
REFERENCE   2  (bases 1 to 3847)
  AUTHORS   Doblin,M.S., De Melis,L., Newbigin,E., Bacic,A. and Read,S.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Plant Biology, University of California, 1
            Shields Ave, Davis, CA 95616, USA
FEATURES             Location/Qualifiers
     source          1. .3847
                     /organism="Nicotiana alata"
                     /mol_type="mRNA"
                     /db_xref="taxon:4087"
                     /tissue_type="pistil"
                     /note="genotype: S2S2"
     gene            1. .3847
                     /gene="CesA1"
     CDS             103. .3378
                     /gene="CesA1"
                     /codon_start=1
                     /product="cellulose synthase catalytic subunit"
                     /protein_id="AAK49454.1"
                     /db_xref="GI:13925881"
                     /translation="MDTKGRLVAGSHNRNEFVVINADDVGRVTSVKELSGQICQICGD
                     EIEVTVDGEPFIACNECAFPVCRQCYEYERREGNQACPQCKTRFKRIKGSPRVDGDDE
                     DDEFDDLDHEFDYHGNPRYMSEAAFSSRLGRGTNHNASGLTTPSEVDPAALNSEIPLL
                     TYGQEDDTISADKHALIIPPFMGRGKKVHPVPYSDSMSLPPRPMDPKKDLAVYGYGTV
                     AWKERMEDWKKKQNDKLQVVKHGGGKGGGNDGDELDDPDLPKMDEGRQPLSRKLPISS
                     SRLSPYRLLILVRLAVVGLFFHYRITHPVNDAYALWLISIICEIWFAVSWIFDQFPKW
                     FPIVRETYLDRLSLRYEKEGKPSGLAPIDIFVSTVDPLKEPPLITANTVLSILAVDYP
                     EDKVSCYVSDDGAAMLTFEALSETSEFARKWVPFCKKFNIEPRAPEWYFSQKVDYLKN
                     KVHPSFVRERRAMKRDYEEFKVRINGLVATAQKVPEDGWTMQDGTPWPGNLVRDHPGM
                     IQVFLGNDGVRDIEGNVLPRLIYVSREKRPGFDHHKKAGAMNALMRVSAVISNAPYLL
                     NVDCDHYINNSKALREAMCFMMDPTSGKKICYVQFPQRFDGIDRHDRYSNRNVVFFDI
                     NMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKKTKPPGKTCNCWPKWCCCCFGSRKKH
                     KKGKTTKDNKKKTKTKEASPQIHALENIEEGIEGIDSEKATLMPQIKLEKKFGQSPVF
                     VASTLLEDGGIPPGATSASLLKEAIHVISCGYEDKTEWGREVGWIYGSVTEDILTGFK
                     MHCHGWRSVYCMPKRPAFKGSAPINLSDRLHQVLRWALGSVEILLSKHCPIWYGYGCG
                     LKPLERFSYINSVVYPLTSLPLIAYCALPAVCLLTGKFIVPEISNYASILFMGLFIMI
                     AATSVLEMQWGGVTIDDWWRNEQFWVIGGASSHLFALFQGLLKVLAGVSTSFTVTSKA
                     ADDGEFSELYLFKWTSLLIPPMTLLIINIIGVIVGISDAINNGYDSWGPLFGRLFFAL
                     WVIVHLYPFLKGVMGRQNKVPTIIVVWSILLASIFSLLWVRVNPFTARGGLVLEVCGL
                     DCE"
BASE COUNT     1061 a    719 c    925 g   1142 t
ORIGIN
```

```
  Query Match             38.9%;  Score 1531.4;  DB 8;  Length 3847;
  Best Local Similarity   68.2%;  Pred. No. 8.8e-286;
  Matches 2252;  Conservative    0;  Mismatches  996;  Indels   53;  Gaps    7;

Qy        139 AAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCG 198
              |||  |||||  | |  |  |||| || || || || |||||  | || ||| | || |
Db         98 AAGTAATGGATACTAAAGGGAGGCTAGTTGCTGGTTCACACAATAGGAATGAGTTTGTTG 157

Qy        199 TCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGT 258
              |||||       || |  | |    |   |   |   ||    |  ||||||  | |
Db        158 TCATCAATGCTGATGATGTTGGAAGAGTAACTTCTGTGAAAGAATTAAGTGGGCAGATTT 217

Qy        259 GCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCA 318
              |||||||||| || || ||  | |  || |   |  ||||| || ||  | || ||||
Db        218 GCCAGATTTGTGGAGATGAGATTGAAGTTACAGTAGATGGGGAGCCATTTATTGCTTGCA 277

Qy        319 ACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGC 378
              |||| || || ||||| || ||| |  | ||||| || || ||  |  | || || |  |
Db        278 ACGAATGTGCATTCCCTGTTTGCAGACAATGCTATGAGTATGAAAGGAGAGAAGGAAATC 337

Qy        379 AGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCG 438
              |    ||||| || ||||| ||||| | ||||||| | || ||  |||   | ||    |
Db        338 AAGCTTGCCCTCAATGCAAAACTCGGTTCAAGCGCATTAAAGGGAGCCCCAGGGTTGATG 397

Qy        439 GTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATG 498
              | || || || || || |   | |||||| | ||| | |||||| | |   | || |
Db        398 GAGATGATGAAGATGATGAATTTGATGACTTAGACCATGAGTTCGATTACCATGG-TAAC 456

Qy        499 ACTCGCAGTCTGTGGCCGAGTCCATGCTCTAC-------GGCCACATGAGCTACGGCCGT 551
               ||||   | |||     |  | ||| |||| |        ||||  |  | |      |
Db        457 CCTCGATATATGTCTGAGGCTGCATTCTCTTCTCGTCTTGGCCGCGGTACCAATCATAAT 516

Qy        552 GGAGGTGACCCTAATGGCGCGCCACAAG------CTTTCCAGCTCAACCCCAATGTTCCA 605
              |    ||  | || |  | |  || |||      ||      || ||| |  |  | ||
Db        517 GCTTCTGGGCTTACTACCCCATCAGAAGTAGATCCTGCTGCTCTTAACTCAGAGATCCCT 576

Qy        606 CTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCT 665
              || || ||  | || |||  ||  |||   ||  |  | ||  |||| || ||  |   |
Db        577 CTTCTTACTTATGGTCAAGAGGATGATACAATTTCCGCTGACAAGCATGCTCTTATCATT 636

Qy        666 TCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTA 725
              |      |    ||| | |||||||   | |||||  | |||||| | ||| |||   |
Db        637 CCACCGTTTATGGGTCGCGGAAAGAAAGTTCATCCTGTGCCTTATTCTGATTCCA---TG 693

Qy        726 CCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGT 785
              || |||  ||  |  | ||||| ||    || ||| | || | |||||  |||| | |
Db        694 TCTTTGCCCCCTCGACCCATGGATCCTAAGAAAGATTTGGCAGTTTATGGATATGGGACT 753

Qy        786 GTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACG 845
              ||||| |||||||| | |||||| | |||||| | | |||  |    |     |||  |
Db        754 GTTGCATGGAAGGAAAGAATGGAGGACTGGAAGAAAAAACAGAATGATAAATTACAGGTG 813

Qy        846 GGGAATGATGGTGGTGGTGATGATGGT--------------------GACGATGCTGAT 884
              |  ||  | || || ||| | | ||||                    || ||| | |||
Db        814 GTTAAGCACGGAGGCGGTAAAGGTGGTGGTAACGATGGAGATGAGCTGGATGATCCCGAT 873

Qy        885 CTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGC 944
               | ||    ||||||||||  |||||| ||| || || ||| | ||  || | || |||
```

```
Db        874 TTGCCCAAGATGGATGAAGGTAGACAACCACTTTCGAGAAAACTGCCTATTTCCTCTAGC 933

Qy        945 CAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTC 1004

Db        934 AGGCTAAGCCCATACAGATTGCTCATTTTAGTTCGTCTTGCTGTCGTTGGGCTCTTTTTC 993

Qy       1005 CACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATC 1064

Db        994 CACTATAGAATAACGCATCCTGTCAATGATGCATATGCGTTGTGGCTGATTTCTATTATA 1053

Qy       1065 TGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATT 1124

Db       1054 TGTGAAATATGGTTTGCGGTATCATGGATATTTGATCAGTTCCCAAAATGGTTTCCGATT 1113

Qy       1125 GAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCT 1184

Db       1114 GTGCGAGAGACATACCTGGACAGGCTATCTCTGAGGTATGAGAAAGAAGGGAAGCCTTCT 1173

Qy       1185 CAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTG 1244

Db       1174 GGGTTAGCCCCTATAGATATATTTGTTAGTACGGTGGATCCCCTGAAAGAACCTCCGCTT 1233

Qy       1245 GTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCT 1304

Db       1234 ATTACTGCAAATACTGTTCTCTCCATTCTCGCTGTGGATTATCCGGAGGACAAGGTTTCT 1293

Qy       1305 TGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCT 1364

Db       1294 TGCTATGTCTCTGACGATGGTGCTGCCATGCTTACTTTTGAGGCTCTCTCTGAAACATCT 1353

Qy       1365 GAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCA 1424

Db       1354 GAGTTCGCAAGGAAATGGGTCCCATTCTGCAAGAAGTTCAACATAGAGCCTCGGGCCCCA 1413

Qy       1425 GAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTT 1484

Db       1414 GAGTGGTATTTCTCTCAGAAGGTTGATTATCTGAAAAATAAAGTACATCCATCATTTGTG 1473

Qy       1485 AGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTA 1544

Db       1474 AGGGAACGTCGTGCTATGAAGAGAGACTATGAAGAATTCAAGGTTCGGATAAATGGACTG 1533

Qy       1545 GTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGG 1604

Db       1534 GTTGCCACGGCACAAAAGGTTCCTGAAGATGGCTGGACCATGCAAGATGGAACGCCTTGG 1593

Qy       1605 CCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGA 1664

Db       1594 CCCGGAAATCTTGTCAGGGATCATCCTGGAATGATCCAGGTCTTCCTGGGTAATGATGGT 1653

Qy       1665 GGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGA 1724

Db       1654 GTCCGTGATATTGAAGGGAATGTACTGCCTCGCCTTATCTATGTTTCTCGTGAGAAGCGT 1713

Qy       1725 CCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCT 1784

Db       1714 CCAGGATTTGACCACCACAAGAAGGCTGGTGCCATGAATGCTTTGATGCGGGTATCAGCG 1773

Qy       1785 GTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGC 1844

Db       1774 GTCATATCAAATGCTCCTTACTTGCTCAATGTGGATTGTGATCACTATATAAATAACAGT 1833

Qy       1845 AAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGC 1904

Db       1834 AAGGCACTGAGAGAAGCTATGTGCTTTATGATGGACCCCACTTCAGGAAAGAAAATATGC 1893

Qy       1905 TATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGG 1964

Db       1894 TATGTACAATTCCCACAAAGGTTTGATGGTATTGACAGGCATGACAGATACTCCAACCGC 1953

Qy       1965 AATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTAT 2024

Db       1954 AACGTGGTCTTCTTTGATATCAATATGAAAGGACTTGACGGGATCCAGGGTCCAATTTAC 2013
```

```
Qy       2025 GTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACA 2084

Db       2014 GTGGGTACTGGATGTGTCTTCAGGAGGCAAGCGCTTTATGGTTATGATGCTCCAAAGAAG 2073

Qy       2085 AAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGC 2144

Db       2074 ACAAAACCTCCAGGAAAAACATGCAACTGCTGGCCAAAATGGTGTTGCTGTTGTTTCGGT 2133

Qy       2145 TTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTA 2204

Db       2134 TCT------------CGAAAGAAGCATAAGAAAGGGAAAACAACCAAGGATAACAAAAAG 2181

Qy       2205 TTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCT 2264

Db       2182 AAGACTAAAACCAAAGAGGCTTCACCACAAATTCACGCCCTTGAAAATATCGAGGAAGGA 2241

Qy       2265 GCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAA 2324

Db       2242 ATTGAAGGAATTGATAGTGAAAAAGCAACCCTCATGCCTCAGATAAAACTTGAGAAGAAA 2301

Qy       2325 TTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAG 2384

Db       2302 TTTGGACAATCACCAGTATTCGTTGCTTCAACACTTCTAGAAGATGGTGGCATTCCTCCA 2361

Qy       2385 AGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAA 2444

Db       2362 GGAGCAACATCAGCATCTCTCCTGAAAGAAGCAATACATGTCATTAGTTGTGGTTATGAA 2421

Qy       2445 GACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATT 2504

Db       2422 GACAAAACAGAATGGGGTAGAGAGGTTGGATGGATTTATGGATCTGTTACTGAGGATATC 2481

Qy       2505 CTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGG 2564

Db       2482 CTTACCGGATTTAAGATGCACTGTCACGGTTGGAGATCTGTCTACTGCATGCCCAAAAGG 2541

Qy       2565 GTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGG 2624

Db       2542 CCTGCGTTTAAGGGGTCAGCTCCTATCAATCTTTCAGATCGTCTTCATCAGGTTCTGCGG 2601

Qy       2625 TGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTAT 2684

Db       2602 TGGGCTTTGGGATCTGTTGAAATTCTCTTGAGTAAACATTGTCCCATTTGGTATGGATAT 2661

Qy       2685 GGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGG 2744

Db       2662 GGGTGTGGGCTGAAGCCGCTGGAGCGCTTTTCCTACATAAACTCAGTGGTCTACCCATTA 2721

Qy       2745 ACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAA 2804

Db       2722 ACTTCTCTTCCTTTGATCGCATATTGTGCATTGCCAGCTGTCTGTTTGCTTACTGGGAAA 2781

Qy       2805 TTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGC 2864

Db       2782 TTTATTGTTCCAGAGATTAGCAACTATGCTAGCATTCTTTTCATGGGACTTTTCATAATG 2841

Qy       2865 ATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGG 2924

Db       2842 ATTGCTGCCACGAGTGTTCTTGAGATGCAATGGGGAGGTGTTACTATTGATGACTGGTGG 2901

Qy       2925 AGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAG 2984

Db       2902 AGAAACGAGCAGTTCTGGGTCATTGGTGGTGCCTCATCTCACCTATTTGCTCTTTTCCAA 2961

Qy       2985 GGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGA 3044

Db       2962 GGTCTTCTTAAAGTTTTGGCTGGTGTCAGCACTAGCTTCACTGTTACTTCCAAAGCAGCA 3021

Qy       3045 GATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCT 3104

Db       3022 GACGACGGGGAGTTTTCTGAGCTTTATCTGTTCAAGTGGACATCTTTGTTGATTCCTCCT 3081
```

```
Qy       3105 ACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAAT 3164
Db       3082 ATGACTTTGTTAATTATTAACATCATTGGAGTCATAGTTGGAATTTCGGATGCAATCAAC 3141

Qy       3165 AACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATT 3224
Db       3142 AACGGTTATGATTCTTGGGGTCCTCTATTCGGTAGGCTTTTCTTCGCCTTGTGGGTCATT 3201

Qy       3225 GTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATT 3284
Db       3202 GTTCATCTATATCCCTTCCTCAAAGGTGTGATGGGCAGACAAAATAAAGTTCCAACCATC 3261

Qy       3285 GTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGAT 3344
Db       3262 ATTGTTGTCTGGTCTATCCTTCTGGCATCTATCTTTTCCTTATTGTGGGTCCGAGTCAAC 3321

Qy       3345 CCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAG 3404
Db       3322 CCATTCACGGCAA---GAGGTGGACTCGTTTTAGAAGTATGCGGGTTGGACTGCGAGTAA 3378

Qy       3405 G 3405
Db       3379 G 3379
```

```
RESULT 5
AX507835
LOCUS       AX507835               3255 bp    DNA     linear   PAT 27-SEP-2002
DEFINITION  Sequence 2530 from Patent WO0216655.
ACCESSION   AX507835
VERSION     AX507835.1  GI:23389072
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 2530 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1. .3255
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
BASE COUNT      841 a    641 c    813 g    960 t
ORIGIN
```

Query Match 38.3%; Score 1509; DB 6; Length 3255;
Best Local Similarity 67.6%; Pred. No. 1.9e-281;
Matches 2207; Conservative 0; Mismatches 1030; Indels 30; Gaps 5;

```
Qy        144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203
Db          1 ATGAACACCGGTGGTCGGTTAATCGCCGGTTCTCACAACAGGAATGAGTTTGTCCTCATT 60

Qy        204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db         61 AATGCCGATGAGAATGCCCGAATAAGATCAGTCCAAGAGCTGAGTGGACAGACATGTCAA 120

Qy        264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db        121 ATCTGCAGAGATGAGATCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAACGAA 180

Qy        324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db        181 TGTGCATTCCCTGTGTGTAGACCTTGCTATGAGTACGAAAGACGAGAAGGCAATCAAGCT 240

Qy        384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
```

```
Db        241 TGTCCACAGTGCAAAACCCGTTTCAAACGTCTTAAAGGAAGTCCAAGAGTTGAAGGTGAT 300

Qy        444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
Db        301 GAAGAGGAAGATGACATTGATGATTTAGACAATGAGTTTGA---GTATGGAAATAATGGG 357

Qy        504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db        358 ATTGGATTTGATCAGGTTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCCA 417

Qy        564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db        418 CAATCTGATTTGGATTCAGCTCCACCTGGCTCTCAGATTCCATTGCTGACTTACGGCGAC 477

Qy        624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGC---TGGTGCCTTCTTTCATGGGTGGT 680
Db        478 GAGGACGTTGAGATTTCTTCTGATAGACATGCTCTTATTGTTCCTCCTTCACTTGGTGGT 537

Qy        681 GGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGG 740
Db        538 CATGGCAATAGAGTTCATCCTGTTTCTCTTTCTGACCCGACCGTGGCTGCACATCCAAGG 597

Qy        741 TCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAA 800
Db        598 CCTATGGTACCTCAGAAAGATCTTGCGGTTTATGGTTATGGAAGTGTCGCTTGGAAAGAT 657

Qy        801 CGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGA---------AT 851
Db        658 CGGATGGAGGAATGGAAGAGAAAGCAGAATGAGAAACTTCAGGTTGTTAGGCATGAAGGA 717

Qy        852 GATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAA 911
Db        718 GATCCTGATTTTGAAGATGGTGATGATGCTGATTTTCCAATGATGGATGAGGGAAGGCAG 777

Qy        912 CAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATC 971
Db        778 CCATTGTCTAGGAAGATACCAATCAAATCGAGCAAGATAAATCCTTACCGGATGTTAATT 837

Qy        972 ATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAAT 1031
Db        838 GTGCTACGTCTTGTGATTCTTGGTCTCTTCTTTCACTACCGTATTCTTCACCCCGTCAAA 897

Qy       1032 GATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGG 1091
Db        898 GATGCATATGCTTTGTGGCTTATTTCTGTTATATGTGAGATATGGTTTGCTGTTTCATGG 957

Qy       1092 ATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTG 1151
Db        958 GTTCTTGATCAGTTCCCTAAATGGTACCCTATCGAGCGAGAAACGTACTTGGACCGACTC 1017

Qy       1152 TCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTC 1211
Db       1018 TCATTAAGATATGAGAAAGAAGGGAAACCGTCGGGACTATCCCCTGTGGATGTATTTGTT 1077

Qy       1212 AGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATC 1271
Db       1078 AGTACAGTGGATCCATTGAAAGAGCCTCCGCTTATTACTGCAAATACTGTCTTGTCTATT 1137

Qy       1272 CTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCA 1331
Db       1138 CTTGCTGTTGATTATCCTGTCGATAAGGTTGCTTGTTACGTATCTGATGATGGTGCTGCT 1197

Qy       1332 ATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTC 1391
Db       1198 ATGCTTACTTTCGAAGCTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTC 1257

Qy       1392 TGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGAC 1451
Db       1258 TGCAAGAAATATTGTATTGAGCCTCGTGCTCCCGAATGGTATTTCTGCCATAAAATGGAC 1317

Qy       1452 TACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAG 1511
```

```
Db      1318 TACTTGAAGAATAAAGTTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGAT 1377

Qy      1512 TATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAA 1571

Db      1378 TATGAAGAATTCAAAGTAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCCTGAG 1437

Qy      1572 GAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCT 1631

Db      1438 GATGGTTGGACTATGCAAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCCT 1497

Qy      1632 GGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTG 1691

Db      1498 GGCATGATTCAGGTCTTCCTTGGAAGTGACGGTGTTCGTGATGTCGAAAACAACGAGTTG 1557

Qy      1692 CCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCT 1751

Db      1558 CCTCGATTAGTTTACGTTTCTCGTGAGAAGAGACCCGGATTTGATCACCATAAGAAGGCT 1617

Qy      1752 GGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTA 1811

Db      1618 GGAGCTATGAATTCCCTGATACGAGTCTCTGGGGTTCTATCAAATGCTCCTTACCTTCTG 1677

Qy      1812 AACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTT 1871

Db      1678 AATGTCGATTGTGATCACTACATCAACAATAGCAAAGCTCTTAGAGAAGCAATGTGTTTC 1737

Qy      1872 ATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGAT 1931

Db      1738 ATGATGGATCCTCAGTCAGGAAAGAAAATCTGTTATGTTCAGTTCCCTCAAAGGTTCGAT 1797

Qy      1932 GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATG 1991

Db      1798 GGGATTGATAGGCACGATCGATACTCAAATCGCAATGTTGTGTTCTTTGATATCAATATG 1857

Qy      1992 AAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGG 2051

Db      1858 AAAGGTTTGGATGGGCTACAAGGGCCTATATACGTCGGTACAGGTTGTGTTTTCAGGAGG 1917

Qy      2052 CAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAAC 2111

Db      1918 CAAGCGCTTTACGGATTTGATGCACCGAAGAAGAAGAAGGGCCCACGTAAGACATGCAAT 1977

Qy      2112 TGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACT 2171

Db      1978 TGCTGGCCAAAATGGTGTCTCCTATGTTTTGGTTCAAGAAAGAATCGTAAAGCAAAGACA 2037

Qy      2172 ACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCC 2231

Db      2038 GTGGCTGCGGATAAGAAGAAGAAGAA----------TAGGGAAGCGTCAAAGCAGATC 2085

Qy      2232 CCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCC 2291

Db      2086 CACGCATTAGAAAATATCGAAGAGGGCCGCGTCACTAAAGGTTCTAACGTAGAACAGTCA 2145

Qy      2292 GGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACA 2351

Db      2146 ACCGAGGCAATGCAAATGAAGTTGGAGAAGAAATTTGGGCAGTCTCCTGTATTTGTTGCA 2205

Qy      2352 TCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAA 2411

Db      2206 TCTGCGCGTATGGAGAATGGTGGGATGGCTAGAAACGCAAGCCCGGCTTGTCTGCTTAAA 2265

Qy      2412 GAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATT 2471

Db      2266 GAAGCCATCCAAGTCATTAGTTGCGGATATGAAGATAAAACTGAATGGGGAAAAGAGATT 2325

Qy      2472 GGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCAT 2531

Db      2326 GGGTGGATCTATGGTTCTGTTACCGAAGATATTCTTACGGGTTTTAAGATGCATTCTCAT 2385

Qy      2532 GGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTG 2591

Db      2386 GGTTGGAGATCTGTTTATTGTACACCAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATC 2445

Qy      2592 AATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651

Db      2446 AATCTTTCGGATCGTCTCCATCAAGTTCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTC 2505

Qy      2652 TTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGA 2711

Db      2506 TTGAGTAGGCATTGTCCTATTTGGTATGGTTATGGAGGTGGGTTGAAATGGCTTGAGCGG 2565

Qy      2712 TTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGT 2771

Db      2566 TTGTCCTACATTAACTCTGTGGTTTACCCGTGGACCTCTCTACCGCTCATCGTTTACTGT 2625

Qy      2772 ACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTT 2831

Db      2626 TCTCTCCCTGCCATCTGTCTTCTCACTGGAAAATTCATCGTTCCCGAGATTAGCAACTAT 2685

Qy      2832 GCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATG 2891

Db      2686 GCGAGTATCCTCTTCATGGCGCTCTTCTCGTCGATTGCAATAACGGGTATTCTCGAGATG 2745

Qy      2892 AGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGA 2951

Db      2746 CAATGGGGCAAAGTTGGGATCGATGATTGGTGGAGAAACGAACAGTTTTGGGTCATTGGA 2805

Qy      2952 GGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTT 3011

Db      2806 GGTGTTTCTGCGCATCTGTTTGCTCTCTTCCAAGGTCTCCTCAAGGTTCTTGCTGGTGTC 2865

Qy      3012 GATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071

Db      2866 GACACTAACTTCACAGTCACATCAAAAGCAGCTGATGATGGAGAGTTCTCTGACCTTTAC 2925

Qy      3072 ACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATT 3131

Db      2926 CTCTTCAAATGGACTTCACTTCTCATCCCTCCAATGACTCTACTCATCATAAACGTCATT 2985

Qy      3132 GGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTC 3191

Db      2986 GGAGTCATAGTCGGAGTCTCTGATGCCATCAGCAATGGATACGACTCGTGGGGACCGCTT 3045

Qy      3192 TTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGT 3251

Db      3046 TTCGGAAGACTGTTCTTTGCACTTTGGGTCATCATTCATCTTTACCCGTTCCTTAAAGGT 3105

Qy      3252 TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCT 3311

Db      3106 TTGCTTGGGAAACAAGATAGAATGCCAACCATTATTGTCGTCTGGTCCATCCTCCTGGCC 3165

Qy      3312 TCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCG 3371

Db      3166 TCGATTCTTACACTTCTTTGGGTCCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCCT 3222

Qy      3372 CTTCTTGAGGAGTGTGGTTTGGATTGC 3398

Db      3223 ATTCTCGAGATCTGTGGTTTAGACTGC 3249
```

RESULT 6
AF062485

LOCUS	AF062485	3444 bp mRNA linear PLN 08-AUG-1998
DEFINITION	Arabidopsis thaliana cellulose synthase mRNA, partial cds.	
ACCESSION	AF062485	
VERSION	AF062485.1 GI:3135610	
KEYWORDS	.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 3444)	
AUTHORS	Wu,L., Joshi,C.P. and Chiang,V.L.	
TITLE	AraxCelA, a new member of the cellulose synthase gene family from	

```
          Arabidopsis (Accession No. AF062485) (PGR98-114)
  JOURNAL Plant Physiol. 117 (3), 1125 (1998)
REFERENCE 2  (bases 1 to 3444)
  AUTHORS Joshi,C.
  TITLE   Direct Submission
  JOURNAL Submitted (01-MAY-1998) Plant Biotechnology Research Center, School
          of Forestry and Wood Products, Michigan Technological University,
          Houghton, MI 49931, USA
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone_lib="ABRC Number CD4-16"
     CDS             <1. .3247
                     /codon_start=2
                     /product="cellulose synthase"
                     /protein_id="AAC29067.1"
                     /db_xref="GI:3135611"
                     /translation="RPRLIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIE
                     LTVDGEPFVACNECAFPVCRPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDD
                     IDDLDNEFEYGNNGIGFDQVSEGMSISRRNSGFPQSDLDSAPPGSQIPLLTYGDEDVE
                     ISSDRHALIVPPSLGGHGNRVHPVSLSDPTVAAHRRLMVPQKDLAVYGYGSVAWKDRM
                     EEWKRKQNEKLQVVRHEGDPDFEDGDDADFPMMDEGRQPLSMKIPIKSSKINPYRMLI
                     VLRLVILGLFFHYRILHPVKDAYALWLISVICEIWFAVSWVLDQFPKWYPIERETYLD
                     RLSLRYEKEGKPSGLSPVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSD
                     DGAAMLTFEALSETAEFARKWVPFCKKYCIEPRAPEWYFCHKMDYLKNKVHPAFVRER
                     RAMKRDYEEFKVKINALVATAQKVPEDGWTMQDGTPWPGNSVRDHPGMIQVFLGSDGV
                     RDVENNELPRLVYVSREKRPGFDHHKKAGAMNSLIRVSGVLSNAPYLLNVDCDHYINN
                     SKALREAMCFMMDPQSGKKICYVQFPQRFDGIDRHDRYSNRNVVFFDINMKGLDGLQG
                     PIYVGTGCVFRRQALYGFDAPKKKKGPRKTCNCWPKWCLLCFGSRKNRKAKTVAADKK
                     KKNREASKQIHALENIEEGRGHKVLNVEQSTEAMQMKLQKKYGQSPVFVASARLENGG
                     MARNASPACLLKEAIQVISRGYEDKTEWGKEIGWIYGSVTEDILTGSKMHSHGWRHVY
                     CTPKLAAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPIWYGYGGGLKWLERLSYI
                     NSVVYPWTSLPLIVYCSLPAICLLTGKFIVPEISNYASILFMALFSSIAITGILEMQW
                     GKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLAGVDTNFTVTSKAADDGEFSDLY
                     LFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYDSWGPLFGRLFFALWVIIHLYPFL
                     KGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVAKGGPILEICGLDCL"
BASE COUNT      903 a    664 c    860 g   1017 t
ORIGIN

  Query Match             37.9%;  Score 1489.8;  DB 8;  Length 3444;
  Best Local Similarity   67.3%;  Pred. No. 9.8e-278;
  Matches 2189;  Conservative    0;  Mismatches 1032;  Indels   30;  Gaps    5;

Qy        160 GGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATC 219
              || |   | ||||| || |||||| | || ||| | ||| ||||       ||| |    ||
Db          9 GGTTAATCGCCGGTTCTCACAACAGGAATGAGTTTGTCCTCATTAATGCCGATGAGAATG 68

Qy        220 CCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACG 279
              || |     |   |     | |||| ||  || |||    || || || ||| | || ||
Db         69 CCCGAATAAGATCAGTCCAAGAGCTGAGTGGACAGACATGTCAAATCTGCAGAGATGAGA 128

Qy        280 TCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCT 339
              |||    |  |     |  || || || || ||||| || ||||| || || ||||| || |
Db        129 TCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAACGAATGTGCATTCCCTGTGT 188

Qy        340 GCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGA 399
              | |    ||||| || ||||||  | || || ||||  ||    || || ||||||||| |
Db        189 GTAGACCTTGCTATGAGTACGAAAGACGAGAAGGCAATCAAGCTTGTCCACAGTGCAAAA 248

Qy        400 CTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCG 459
              | || | ||| || || || ||  | | | | ||    ||||| || ||||| || | |
Db        249 CCCGTTTCAAACGTCTTAAAGGAAGTCCAAGAGTTGAAGGTGATGAAGAGGAAGATGACA 308

Qy        460 TCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGT 519
              | |||||  | ||||| |||||  |   | | ||  || |   |       | |   ||
Db        309 TTGATGATTTAGACAATGAGTTTGA---GTATGGAAATAATGGGATTGGATTTGATCAGG 365

Qy        520 CCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAG 579
                                  ||       ||      ||| ||| |     |       |
Db        366 TTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCCACAATCTGATTTGGATT 425

Qy        580 CTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCC 639
              |   |  |    | |  |  ||||| | || ||  ||||  |   ||  | ||| ||
Db        426 CAGCTCCACCTGGCTCTCAGATTCCATTGCTGACTTACGGCGACGAGGACGTTGAGATTT 485

Qy        640 CACCGGAGCAGCACGCGC---TGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATAC 696
              | | ||    || || |   | || ||| |||   | ||||||   || || ||   | |
Db        486 CTTCTGATAGACATGCTCTTATTGTTCCTCCTTCACTTGGTGGTCATGGCAATAGAGTTC 545

Qy        697 ATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCA 756
              ||||  || ||  | | || || | | |  |||  || |  |||   ||||  ||    |
Db        546 ATCCTGTTTCTCTTTCTGACCCGACCGTGGCTGCACATCGAAGGCTGATGGTACCTCAGA 605

Qy        757 AGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGA 816
              | |||||||| |  ||||| ||||| ||||| |||||||| || |||||||||| | ||||
Db        606 AAGATCTTGCGGTTTATGGTTATGGAAGTGTCGCTTGGAAAGATCGGATGGAGGAATGGA 665

Qy        817 AGCAGAGACAAGAGAGGATGCACCAGACGGGGA---------ATGATGGTGGTGGTGATG 867
              ||   | ||  |   ||  |  |||   |  |         |||  || |  ||| |
Db        666 AGAGAAAGCAGAATGAGAAACTTCAGGTTGTTAGGCATGAAGGAGATCCTGATTTTGAAG 725

Qy        868 ATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAA 927
              ||||||| ||||||||| | ||| | |||||||| | ||| || | | |||| | ||| |
Db        726 ATGGTGATGATGCTGATTTTCCAATGATGGATGAGGGAAGGCAGCCATTGTCTATGAAGA 785

Qy        928 TTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGG 987
              | ||| |   ||| ||| |||| ||||| ||  ||||| | ||  |  | || ||||||
Db        786 TACCAATCAAATCGAGCAAGATAAATCCTTACCGGATGTTAATTGTGCTACGTCTTGTGA 845

Qy        988 TTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGT 1047
              || | ||  |||||| |||||||  |  | || || || || ||||||| ||||||||
Db        846 TTCTTGGTCTCTTCTTTCACTACCGTATTCTTCACCCCGTCAAAGATGCATATGCTTTGT 905

Qy       1048 GGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCC 1107
              |||| || |||||||| |||| || |||||||||  | || ||| ||||||||| ||||
Db        906 GGCTTATTTCTGTTATATGTGAGATATGGTTTGCTGTTTCATGGGTTCTTGATCAGTTCC 965

Qy       1108 CAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACA 1167
              | || |||| |||||| ||| |||| || ||| | ||||| || ||| | || |  || |
Db        966 CTAAATGGTACCCTATCGAGCGAGAAACGTACTTGGACCGACTCTCATTAAGATATGAGA 1025

Qy       1168 AGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCT 1227
              | |||||  | || ||   ||| | ||  | ||| | ||||| ||||| || ||||| |
Db       1026 AAGAAGGGAAACCGTCGGGACTATCCCCTGTGGATGTATTTGTTAGTACAGTGGATCCAT 1085

Qy       1228 TAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATC 1287
              | || || |||||  |  | ||  |||||||||| | ||||| ||| | || |||||||
Db       1086 TGAAAGAGCCTCCGCTTATTACTGCAAATACTGTCTTGTCTATTCTTGCTGTTGATTATC 1145

Qy       1288 CTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAG 1347
              |||| |||||||| |||| || || |||||||||||||||| ||||| || || ||||
Db       1146 CTGTCGATAAGGTTGCTTGTTACGTATCTGATGATGGTGCTGCTATGCTTACTTTCGAAG 1205

Qy       1348 CATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATA 1407
              |  | ||||| ||  ||||||| |||| |||||||||||||||||||||    ||   ||
Db       1206 CTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTCTGCAAGAAATATTGTA 1265

Qy       1408 TTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGG 1467
              |||| ||||| ||||| || ||||| |||    || || || |||||||||| | ||| |
Db       1266 TTGAGCCTCGTGCTCCCGAATGGTATTTCTGCCATAAAATGGACTACTTGAAGAATAAAG 1325

Qy       1468 TGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGG 1527
              |    |  |||||||||||| || |||| |||||||||| ||||| |||||||| |
Db       1326 TTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGATTATGAAGAATTCAAAG 1385

Qy       1528 TGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGC 1587
              | | |||||||| ||||| || | ||| |||||||| ||||| || || ||||| ||||
Db       1386 TAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCCTGAGGATGGTTGGACTATGC 1445

Qy       1588 AAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCT 1647
              |||| || || || ||||| || || | ||| || |||||||||| ||||||||||||
Db       1446 AAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCCTGGCATGATTCAGGTCT 1505
```

```
Qy       1648 TCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATG 1707
             ||||||| ||| | | ||| || || |||| ||||| |||||
Db       1506 TCCTTGGAAGTGACGGTGTTCGTGATGTCGAAAACAACGAGTTGCCTCGATTAGTTTACG 1565

Qy       1708 TTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCAT 1767
             ||||| | ||||| |||| || | | | || |||||||| ||||| |||||||||| |
Db       1566 TTTCTCGTGAGAAGAGACCCGGATTTGATCACCATAAGAAGGCTGGAGCTATGAATTCCC 1625

Qy       1768 TGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATC 1827
             || | |||||||||||  || ||| |||||||||| || |  | ||  | ||||||||||
Db       1626 TGATACGAGTCTCTGGGGTTCTATCAAATGCTCCTTACCTTCTGAATGTCGATTGTGATC 1685

Qy       1828 ACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTAC 1887
             |||||||||||| ||||| ||| | |  ||||||||||||||| |||||||| |||
Db       1686 ACTACATCAACAATAGCAAAGCTCTTAGAGAAGCAATGTGTTTCATGATGGATCCTCAGT 1745

Qy       1888 TAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATG 1947
             |||||||||  | || ||||| |||||||||||||| || ||||||||||||| | || |
Db       1746 CAGGAAAGAAAATCTGTTATGTTCAGTTCCCTCAAAGGTTCGATGGGATTGATAGGCACG 1805

Qy       1948 ACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTA 2007
             | |||||  | || || |||||||| || |||||||||| ||||||||||||||||||
Db       1806 ATCGATACTCAAATCGCAATGTTGTGTTCTTTGATATCAATATGAAAGGTTTGGATGGGC 1865

Qy       2008 TTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTT 2067
             | || || || || || || ||||| || ||||| || || ||||| ||  | || || |
Db       1866 TACAAGGGCCTATATACGTCGGTACAGGTTGTGTTTTCAGGAGGCAAGCGCTTTACGGAT 1925

Qy       2068 ATGATGCCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGT 2127
             |||||| || || |  ||||||  |||  | ||| ||||| |||||||| || ||||
Db       1926 TTGATGCACCGAAGAAGAAGAAGGGCCCACGTAAGACATGCAATTGCTGGCCAAAATGGT 1985

Qy       2128 GCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAG 2187
             |  |    ||||  | |  | || |    ||  |||||          | |
Db       1986 GTCTCCTATGTTTTGGTTCAAGAAAGAATCGTAAAGCAAAGACAGTGGCTGCGGATAAGA 2045

Qy       2188 AGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTG 2247
             ||||||| ||                | | |||     || ||   || ||||  |    |
Db       2046 AGAAGAAGAA------------TAGGGAAGCGTCAAAGCAGATCCACGCATTAGAAAATA 2093

Qy       2248 GTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAAC 2307
                |||  | |   | |  |  ||     | |    ||| || |         | ||  |||
Db       2094 TCGAAGAGGGCCGCGGTCATAAAGTTCTTAACGTAGAACAGTCAACCGAGGCAATGCAAA 2153

Qy       2308 AAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGA 2367
                || ||  | ||||||| ||| || ||| |||| |||||| ||||  | | ||| ||||
Db       2154 TGAAGTTGCAGAAGAAATATGGGCAGTCTCCTGTATTTGTTGCATCTGCGCGTCTGGAGA 2213

Qy       2368 ATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCA 2427
             |||||||  |     |  ||||| || |||| |||  | |||||||| || || ||||
Db       2214 ATGGTGGGATGGCTAGAAACGCAAGCCCGGCTTGTCTGCTTAAAGAAGCCATCCAAGTCA 2273

Qy       2428 TTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGAT 2487
             |||||  | || |||||||| || || || |||||||||||||||  ||||||||||| |
Db       2274 TTAGTCGCGGATATGAAGATAAAACTGAATGGGGAAAAGAGATTGGGTGGATCTATGGTT 2333

Qy       2488 CAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTT 2547
             | ||||| |||||||||| || ||||  |||||||||| |||||||||  |     |||
Db       2334 CTGTTACCGAAGATATTCTTACGGGTTCTAAGATGCATTCTCATGGTTGGAGACATGTTT 2393

Qy       2548 ACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTC 2607
             | || | ||| ||    |  || |||||||| || || ||  | |||||||| |||||||
Db       2394 ATTGTACACCAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATCAATCTTTCGGATCGTC 2453

Qy       2608 TTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCC 2667
             | || || || ||||| ||||| |||||||| ||||||| ||||| || |  |||||  |
Db       2454 TCCATCAAGTTCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTCTTGAGTAGGCATTGTC 2513

Qy       2668 CTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACT 2727
             || |||||||||| ||||| || ||  ||||||   | ||  | || |||||||| ||||
Db       2514 CTATTTGGTATGGTTATGGAGGTGGGTTGAAATGGCTTGAGCGGTTGTCCTACATTAACT 2573

Qy       2728 CCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCT 2787
```

```
                |  | || || || ||||| ||| | || ||| | | ||||||| |  | ||||||||||
Db       2574 CTGTGGTTTACCCGTGGACCTCTCTACCGCTCATCGTTTACTGTTCTCTCCCTGCCATCT 2633

Qy       2788 GTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCA 2847
             || |  | || || ||||| |||  ||| ||| | |  ||  ||| ||  |    ||||
Db       2634 GTCTTCTCACTGGAAAATTCATCGTTCCCGAGATTAGCAACTATGCGAGTATCCTCTTCA 2693

Qy       2848 TGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTG 2907
             || | || ||    | |||      ||| | || || || |||  |||| |    ||||
Db       2694 TGGCGCTCTTCTCGTCGATTGCAATAACGGGTATTCTCGAGATGCAATGGGGCAAAGTTG 2753

Qy       2908 GAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACC 2967
             | || ||||| |||||||| || || ||||| ||||||||||||||||| ||  | || |
Db       2754 GGATCGATGATTGGTGGAGAAACGAACAGTTTTGGGTCATTGGAGGTGTTTCTGCGCATC 2813

Qy       2968 TCTTTGCTGTGTTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCG 3027
             | |||||| | ||||| || || |||||||  | |||||||| || || | |||||| |
Db       2814 TGTTTGCTCTCTTCCAAGGTCTCCTCAAGGTTCTTGCTGGTGTCGACACTAACTTCACAG 2873

Qy       3028 TGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTA 3087
             | |||||||| |  | ||||||| |||||||| || || ||     ||||||||||||
Db       2874 TCACATCAAAAGCAGCTGATGATGGAGAGTTCTCTGACCTTTACCTCTTCAAATGGACTT 2933

Qy       3088 CCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCG 3147
             |  |  | || ||||| |  ||  | ||  |  | ||| ||||||| ||  |  |  || |
Db       2934 CACTTCTCATCCCTCCAATGACTCTACTCATCATAAACGTCATTGGAGTCATAGTCGGAG 2993

Qy       3148 TTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCT 3207
             | ||  |||| |||| || ||||| || || ||||| || || || || ||  || ||||
Db       2994 TCTCTGATGCCATCAGCAATGGATACGACTCGTGGGGACCGCTTTTCGGAAGACTGTTCT 3053

Qy       3208 TTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAA 3267
             ||||| ||||||| ||  | |||||||| || || || |||||||||  |||| |  |||
Db       3054 TTGCACTTTGGGTCATCATTCATCTTTACCCGTTCCTTAAAGGTTTGCTTGGGAAACAAG 3113

Qy       3268 ACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCC 3327
             | || |  ||||| ||| |  |||||||||||| || ||||| || ||   |  | || |
Db       3114 ATAGAATGCCAACCATTATTGTCGTCTGGTCCATCCTCCTGGCCTCGATTCTTACACTTC 3173

Qy       3328 TTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTG 3387
             ||||||| ||| || |||| ||  | ||||   | |  |||||  |||| |||   ||||
Db       3174 TTTGGGTCCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCCTATTCTCGAGATCTGTG 3230

Qy       3388 GTTTGGATTGC 3398
                 |||| || |||
Db       3231 GTTTAGACTGC 3241
```

RESULT 7
AF027173
LOCUS AF027173 3771 bp mRNA linear PLN 07-FEB-1998
DEFINITION Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-A) mRNA, complete cds.
ACCESSION AF027173
VERSION AF027173.1 GI:2827140
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 3771)
AUTHORS Arioli,T., Peng,L., Betzner,A.S., Burn,J., Wittke,W., Herth,W., Camilleri,C., Hofte,H., Plazinski,J., Birch,R., Cork,A., Glover,J., Redmond,J. and Williamson,R.E.
TITLE Molecular analysis of cellulose biosynthesis in Arabidopsis
JOURNAL Science 279 (5351), 717-720 (1998)
MEDLINE 98111412
PUBMED 9445479
REFERENCE 2 (bases 1 to 3771)
AUTHORS Arioli,T.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1997) Plant Science Centre, Australian National

```
          University, Acton, Canberra, ACT 200, Australia
FEATURES             Location/Qualifiers
     source          1. .3771
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
     gene            1. .3771
                     /gene="Ath-A"
     CDS             217. .3471
                     /gene="Ath-A"
                     /note="RSW1-like"
                     /codon_start=1
                     /product="cellulose synthase catalytic subunit"
                     /protein_id="AAC39335.1"
                     /db_xref="GI:2827141"
                     /translation="MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGD
                     EIELTVSSELFVACNECAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDE
                     EEEDIDDLEYEFDHGMDPEHAAEAALSSRLNTGRGGLDSAPPGSQIPLLTYCDEDADM
                     YSDRHALIVPPSTGYGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEV
                     WKRRQGEKLQVIKHEGGNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPY
                     RMLILCRLAILGLFFHYRILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERE
                     TYLDRLSLRYEKEGKPSGLAPVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVAC
                     YVSDDGAAMLTFEALSDTAEFARKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAF
                     VRERRAMKRDYEEFKVKINALVATAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLG
                     HSGVRDTDGNELPRLVYVSREKRPGFDHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDH
                     YINNSKAIRESMCFMMDPQSGKKVCYVQFPQRFDGIDRHDRYSNRNVVFFDINMKGLD
                     GIQGPIYVGTGCVFRRQALYGFDAPKKKKPPGKTCNCWPKWCCLCCGLRKKSKTKAKD
                     KKTNTKETSKQIHALENVDEGVIVPVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQN
                     GGVPRNASPACLLREAIQVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHCHGWRS
                     VYCMPKRAAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPIWYGYGGGLKWLERFS
                     YINSVVYPWTSLPLIVYCSLPAVCLLTGKFIVPEISNYAGILFMLMFISIAVTGILEM
                     QWGGVGIDDWWRNEQFWVIGGASSHLFALFQGLLKVLAGVNTNFTVTSKAADDGAFSE
                     LYIFKWTTLLIPPTTLLIINIIGVIVGVSDAISNGYDSWGPLFGRLFFALWVIVHLYP
                     FLKGMLGKQDKMPTIIVVWSILLASILTLLWVRVNPFVAKGGPVLEICGLNCGN"
BASE COUNT     1037 a    699 c    907 g   1128 t
ORIGIN

  Query Match             36.4%;  Score 1434.2;  DB 8;  Length 3771;
  Best Local Similarity   65.9%;  Pred. No. 5.7e-267;
  Matches 2175;  Conservative    0;  Mismatches 1063;  Indels   63;  Gaps    4;

Qy        124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
              ||   || |  || ||     ||| |  |  | |   ||||  | || |||||  ||||||
Db        197 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 256

Qy        184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
              | |||||  ||||  ||||   |  ||| |    | || |       |    | || |
Db        257 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 316

Qy        244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
              ||  |||||    || || || || || || ||  |||   |  |     || | || |
Db        317 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 376

Qy        304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
              |||  || || |||||||| ||||| ||||| || ||   |    |||||  || || || |
Db        377 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 436

Qy        364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
              |  | || || |  ||     || || |||||||| ||||||||||  |  | || ||
Db        437 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 496

Qy        424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
              | | ||| ||    || || || || || || |  | |  |||       |  |
Db        497 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 556

Qy        484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
                  | ||    ||| |  |   ||  || ||  ||   ||||     | | | | |
Db        557 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 616

Qy        544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
              |||  ||||| ||       |   |  | ||     ||                |  |||
Db        617 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT--------------CAGATTC 661

Qy        604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
              | ||  | ||  |  |  |    |  | ||| ||     | ||  |  ||  |  |
Db        662 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 721

Qy        664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
                ||          ||    || ||  |  | ||||    || | | | ||| |
Db        722 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 781

Qy        724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
              ||||   ||  | || || |||| ||    || ||| |||| | ||||||  ||||| |
Db        782 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 841

Qy        784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
              |||||||||||||||| || |||||   |||||||    ||||||     | ||  ||
Db        842 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 901

Qy        844 CGGGGAATGATGGTGGTGGTGATGATGGT----------------------------GACG 876
                    ||  |||  || ||   | |||||                            || |
Db        902 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 961

Qy        877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
              || |||  | ||  | ||||||||| |||||||  || || || ||  | ||  |||
Db        962 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1021

Qy        937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
              ||||||   || ||||| || |||||| | ||  |   |||  || | | || | ||
Db       1022 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1081

Qy        997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
              | ||||| || ||  || |  | ||||| || |||||||||| ||  || ||| | |  |
Db       1082 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1141

Qy       1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
              | ||||| || || || |||||||| |||||||||||||||||||||| || ||||
Db       1142 CAGTTATATGCGAGATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1201

Qy       1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
              |||||  ||  | || || |||||  ||  | || || || |||| ||| ||||||||
Db       1202 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1261

Qy       1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
              | ||  ||   | | || ||  ||||| | ||||| ||||| || ||||| || || || |
Db       1262 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1321

Qy       1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
              | || ||| | ||| |||| || ||||| || || ||  | || ||||||||||| ||||
Db       1322 CACCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1381

Qy       1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
              ||||| | || ||||| || || |||||||| || ||||| || |||||||||  | ||||
Db       1382 AGGTTGCGTGTTATGTATCAGACGATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1441

Qy       1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
              | ||| |||| ||||| | |||||||||||||| || ||   ||  ||||| || || |
Db       1442 ATACAGCTGAGTTTGCTAGAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1501

Qy       1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
              | ||||| |||||||| ||     |||||||| || ||| ||||  |||| ||      |
Db       1502 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1561

Qy       1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
              |||||  |||||  |  | || |||||||||| |||||||| || || ||||  || |
Db       1562 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAGGAGTTTAAAGTGAAGATAA 1621

Qy       1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
              ||||  | |||||  |  || |||||||| ||||| |||||  ||||| ||||||||||||
Db       1622 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAAGGTTGGACTATGCAAGATGGAA 1681

Qy       1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
              | || |||||||||||||| || ||||| |||||||||||||||||||| ||| | || |
Db       1682 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1741
```

```
Qy       1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db       1742 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1801

Qy       1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db       1802 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1861

Qy       1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
Db       1862 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1921

Qy       1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db       1922 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 1981

Qy       1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db       1982 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2041

Qy       1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db       2042 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2101

Qy       2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db       2102 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAGACAGGCTCTTTATGGTTTTGATGCAC 2161

Qy       2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db       2162 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2221

Qy       2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db       2222 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCAAAGATAAGAAAACTAACACTAAAG 2281

Qy       2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db       2282 AGACTTCA------------------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2323

Qy       2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db       2324 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2383

Qy       2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db       2384 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2443

Qy       2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db       2444 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2503

Qy       2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db       2504 GGTACGAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2563

Qy       2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db       2564 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2623

Qy       2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db       2624 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2683

Qy       2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db       2684 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2743

Qy       2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
Db       2744 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2803

Qy       2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796

Db       2804 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2863

Qy       2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
Db       2864 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2923

Qy       2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
Db       2924 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 2983

Qy       2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
Db       2984 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3043

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
Db       3044 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3103

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db       3104 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3163

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db       3164 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3223

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db       3224 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3283

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db       3284 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3343

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db       3344 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3403

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
Db       3404 GAGTTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3460

Qy       3397 G 3397
              |
Db       3461 G 3461
```

```
RESULT 8
AX505864
LOCUS       AX505864               3255 bp    DNA     linear   PAT 27-SEP-2002
DEFINITION  Sequence 559 from Patent WO0216655.
ACCESSION   AX505864
VERSION     AX505864.1  GI:23387101
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 559 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1. .3255
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
BASE COUNT      884 a    621 c    798 g    952 t
```

ORIGIN

Query Match 36.4%; Score 1431.8; DB 6; Length 3255;
Best Local Similarity 66.0%; Pred. No. 1.6e-266;
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

```
Qy      144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203
Db        1 ATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACAGAAACGAATTCGTTCTCATT 60

Qy      204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db       61 AACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAACTGAGTGGGCAAACATGTCAA 120

Qy      264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db      121 ATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGCTCTTTGTTGCTTGCAACGAA 180

Qy      324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db      181 TGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAACGTAGAGAAGGAAATCAAGCT 240

Qy      384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
Db      241 TGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTAGTCCACGGGTTGATGGAGAT 300

Qy      444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
Db      301 GATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGGATGGACCCT 360

Qy      504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db      361 GAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACACCGGTCGTGGTGGATTGGAT 420

Qy      564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db      421 TCAGCTCCACCTG----------------GCTCTCAGATTCCTCTTTTGACTTATTGTGAT 465

Qy      624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGG 683
Db      466 GAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCGTGCCTCCTTCAACGGGATAT 525

Qy      684 GGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCT 743
Db      526 GGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTGCACCTCCACAGGCGAGATCA 585

Qy      744 ATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGG 803
Db      586 ATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAAGTGTTGCTTGGAAGGACCGT 645

Qy      804 ATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGT 863
Db      646 ATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAGTCATTAAGCATGAAGGAGGA 705

Qy      864 GATGATGGT----------------------------GACGATGCTGATCTACCACTAATG 896
Db      706 AACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATGATCCTGACATGCCTATGATG 765

Qy      897 GATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCA 956
Db      766 GATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTCGTTCAAGCAGAATAAATCCT 825

Qy      957 TATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTG 1016
Db      826 TACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTCTTTTCTTTCATTATAGAATT 885

Qy     1017 ATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGG 1076
Db      886 CTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGTCAGTTATATGCGAGATATGG 945

Qy     1077 TTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACT 1136
Db      946 TTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGTATCCTATAGAACGTGAAACA 1005

Qy     1137 TACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCA 1196
Db      1006 TACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAAAACCGTCAGGATTAGCACCT 1065

Qy     1197 ATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAAT 1256
Db      1066 GTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGCCACCCTTGATTACAGCAAAC 1125

Qy     1257 ACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCT 1316
Db      1126 ACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATAAGGTTGCGTGTTATGTATCA 1185

Qy     1317 GATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAG 1376
Db      1186 GACGATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTGATACAGCTGAGTTTGCTAGA 1245

Qy     1377 AAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTC 1436
Db      1246 AAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCACGAGCTCCTGAGTGGTATTTT 1305

Qy     1437 CAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGA 1496
Db      1306 TCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTGCTTTTGTCAGGGAACGTCGT 1365

Qy     1497 GCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCC 1556
Db      1366 GCTATGAAGAGAGATTATGAGGAGTTTAAAGTGAAGATAAATGCACTGGTTGCTACTGCA 1425

Qy     1557 CAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAAT 1616
Db      1426 CAGAAAGTGCCTGAGGAAGGTTGGACTATGCAAGATGGAACTCCTTGGCCTGGAAACAAC 1485

Qy     1617 GTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGT 1676
Db      1486 GTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTCATAGTGGAGTTCGTGATACG 1545

Qy     1677 GAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAAC 1736
Db      1546 GATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTGAGAAGCGGCCTGGATTTGAT 1605

Qy     1737 CATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAAT 1796
Db      1606 CACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAGTCTCTGCTGTTCTATCAAAC 1665

Qy     1797 GCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAG 1856
Db      1666 GCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCAACAACAGCAAAGCAATTAGA 1725

Qy     1857 GAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTC 1916
Db      1726 GAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGAAAGTTTGTTATGTTCAGTTT 1785

Qy     1917 CCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTT 1976
Db      1786 CCGCAGAGATTTGATGGGATTGATAGACATGATAGATACTCAAACCGTAACGTTGTGTTC 1845

Qy     1977 TTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGA 2036
Db      1846 TTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGACCGATATATGTCGGGACAGGT 1905

Qy     2037 TGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCA 2096
Db      1906 TGTGTGTTTAGAAGACAGGCTCTTTATGGTTTTGATGCACCAAAGAAGAAGAAACCACCA 1965

Qy     2097 TCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGG 2156
Db      1966 GGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGTGTTGTGGGTTGAGAAAGAAG 2025

Qy     2157 AAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAA 2216
Db      2026 AGTAAAACGAAAGCCAAAGATAAGAAAACTAACACTAAAGAGACTTCA---------- 2073
```

```
Qy       2217 GAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCT 2276
Db       2074 ------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAGGTGTTATCGTCCCAGTGTCA 2127

Qy       2277 GAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCT 2336
Db       2128 AATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGGAGAAGAAGTTTGGACAATCT 2187

Qy       2337 TCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCT 2396
Db       2188 CCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAGTTCCCCGTAACGCAAGCCCC 2247

Qy       2397 GCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGAC 2456
Db       2248 GCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGAA 2307

Qy       2457 TGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTC 2516
Db       2308 TGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTGAAGATATCCTGACGGGTTTC 2367

Qy       2517 AAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAA 2576
Db       2368 AAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGCCTAAGCGTGCAGCTTTTAAA 2427

Qy       2577 GGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGG 2636
Db       2428 GGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAGTTCTACGTTGGGCTCTTGGC 2487

Qy       2637 TCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTG 2696
Db       2488 TCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGTATGGTTATGGTGGTGGTTTA 2547

Qy       2697 AAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCC 2756
Db       2548 AAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCTATCCTTGGACTTCACTTCCA 2607

Qy       2757 CTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCA 2816
Db       2608 TTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCACAGGAAAATTCATCGTCCCT 2667

Qy       2817 GAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACG 2876
Db       2668 GAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGTTCATATCCATAGCAGTAACT 2727

Qy       2877 AGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAG 2936
Db       2728 GGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATGATTGGTGGAGAAACGAGCAG 2787

Qy       2937 TTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAG 2996
Db       2788 TTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTCTGTTTCAAGGTTTGCTCAAA 2847

Qy       2997 GTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAG 3056
Db       2848 GTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAAAAGCAGCAGACGATGGAGCT 2907

Qy       3057 TTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTT 3116
Db       2908 TTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGATTCCTCCGACAACACTTCTG 2967

Qy       3117 CTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAG 3176
Db       2968 ATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATGCCATTAGCAATGGCTATGAC 3027

Qy       3177 TCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTAT 3236
Db       3028 TCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTTGGGTCATTGTTCATTTATAC 3087

Qy       3237 CCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGG 3296
Db       3088 CCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGCCTACGATTATTGTGGTCTGG 3147

Qy       3297 TCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCG 3356
```

```
Db       3148 TCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCAGAGTTAACCCGTTTGTGGCT 3207

Qy       3357 AAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTG 3397
Db       3208 A---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATTG 3245


RESULT 9
AR267558
LOCUS       AR267558               3828 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 7 from patent US 6495740.
ACCESSION   AR267558
VERSION     AR267558.1  GI:29697634
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3828)
  AUTHORS   Arioli,A., Williamson,R.E., Betzner,A.S. and Peng,L.
  TITLE     Manipulation of cellulose and/or .beta.-1,4-Glucan
  JOURNAL   Patent: US 6495740-A 7 17-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .3828
                     /organism="unknown"
BASE COUNT     1056 a    721 c    912 g   1139 t
ORIGIN

  Query Match            36.3%;  Score 1429.4;  DB 6;  Length 3828;
  Best Local Similarity  65.8%;  Pred. No. 4.8e-266;
  Matches 2172;  Conservative    0;  Mismatches 1066;  Indels   63;  Gaps    4;

Qy       124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
Db       219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy       184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
Db       279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy       244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
Db       339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy       304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
Db       399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy       364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
Db       459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy       424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
Db       519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy       484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
Db       579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638

Qy       544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
Db       639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT---------------CAGATTC 683

Qy       604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
Db       684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743

Qy       664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
Db       744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803

Qy       724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
```

```
Db        804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863
Qy        784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
Db        864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923
Qy        844 CGGGGAATGATGGTGGTGGTGATGATGGT----------------------------GACG 876
Db        924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983
Qy        877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
Db        984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043
Qy        937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
Db       1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103
Qy        997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
Db       1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163
Qy       1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
Db       1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223
Qy       1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
Db       1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283
Qy       1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
Db       1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343
Qy       1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
Db       1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403
Qy       1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
Db       1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463
Qy       1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
Db       1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523
Qy       1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
Db       1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583
Qy       1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
Db       1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643
Qy       1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
Db       1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703
Qy       1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
Db       1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763
Qy       1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db       1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823
Qy       1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db       1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883
Qy       1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
```

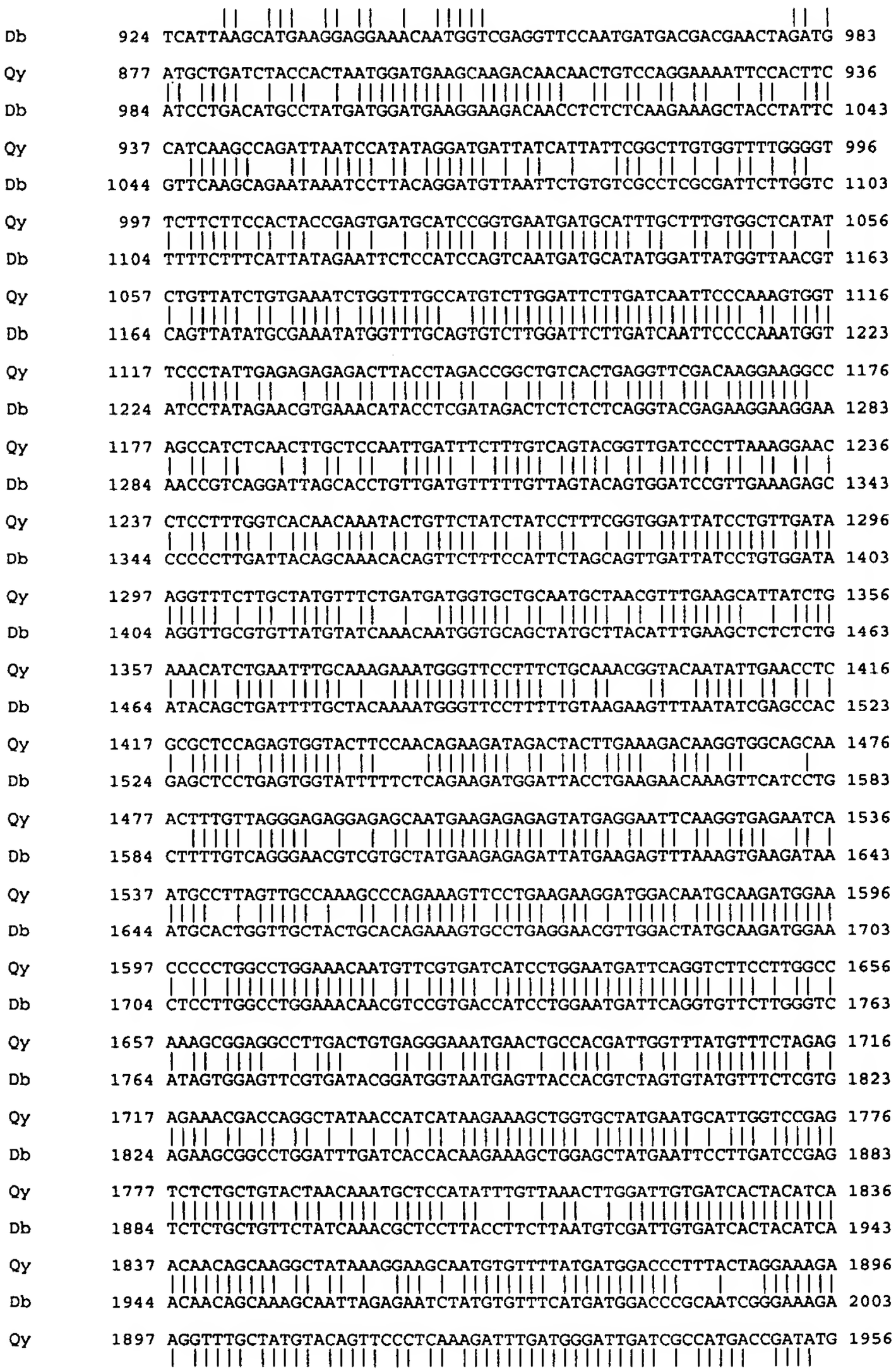
```
Db       1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943
Qy       1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db       1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003
Qy       1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db       2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063
Qy       1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db       2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123
Qy       2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db       2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183
Qy       2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db       2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243
Qy       2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db       2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303
Qy       2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db       2304 AGACTTCA------------------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345
Qy       2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db       2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405
Qy       2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db       2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465
Qy       2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db       2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525
Qy       2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db       2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585
Qy       2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db       2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645
Qy       2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db       2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705
Qy       2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db       2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765
Qy       2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
Db       2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825
Qy       2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796
Db       2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885
Qy       2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
Db       2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945
Qy       2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
Db       2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005
```

```
Qy       2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
Db       3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
Db       3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db       3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db       3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db       3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db       3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db       3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397
              |
Db       3483 G 3483
```

```
RESULT 10
AX030944
LOCUS       AX030944                3828 bp    DNA     linear   PAT 24-NOV-2000
DEFINITION  Sequence 7 from Patent WO9800549.
ACCESSION   AX030944 AX030958
VERSION     AX030944.1  GI:10278347
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Williamson,R.E., Peng,L., Arioli,A. and Betzner,A.S.
  TITLE     Manipulation of cellulose and/or beta -1,4-glucan
  JOURNAL   Patent: WO 9800549-A 7 08-JAN-1998;
            WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI
            ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;
            COMMW SCIENT IND RES ORG (AU)
COMMENT     On Oct 15, 2002 this sequence version replaced gi:10278361.
FEATURES             Location/Qualifiers
     source          1. .3828
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="COLUMBIA"
                     /db_xref="taxon:3702"
                     /clone="ATH-A"
     CDS             239. .3493
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09626.1"
                     /db_xref="GI:10278348"
                     /db_xref="REMTREMBL:CAC09626"
                     /translation="MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGD
                     EIELTVSSELFVACNECAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDE
                     EEEDIDDLEYEFDHGMDPEHAAEAALSSRLNTGRGGLDSAPPGSQIPLLTYCDEDADM
                     YSDRHALIVPPSTGYGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEV
                     WKRRQGEKLQVIKHEGGNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPY
                     RMLILCRLAILGLFFHYRILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERE
                     TYLDRLSLRYEKEGKPSGLAPVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVAC
                     YVSNNGAAMLTFEALSDTADFATKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAF
                     VRERRAMKRDYEEFKVKINALVATAQKVPEERWTMQDGTPWPGNNVRDHPGMIQVFLG
                     HSGVRDTDGNELPRLVYVSREKRPGFDHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDH
                     YINNSKAIRESMCFMMDPQSGKKVCYVQFPQRFDGIDRHDRYSNRNVVFFDINMKGLD
                     GIQGPIYVGTGCVFRKQALYGFDAPKKKKPPGKTCNCWPKWCCLCCGLRKKSKTKATD
                     KKTNTKETSKQIHALENVDEGVIVPVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQN
                     GGVPRNASPACLLREAIQVISCGYQDKTEWGKEIGWIYGSVTEDILTGFKMHCHGWRS
                     VYCMPKRAAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPIWYGYGGGLKWLERFS
                     YINSVVYPWTSLPLIVYCSLPAVCLLTGKFIVPEISNYAGILFMLMFISIAVTGILEM
                     QWGGVGIDDWWRNEQFWVIGGASSHLFALFQGLLKVLAGVNTNFTVTSKAADDGAFSE
                     LYIFKWTTLLIPPTTLLIINIIGVIVGVSDAISNGYDSWGPLFGRLFFALWVIVHLYP
                     FLKGMLGKQDKMPTIIVVWSILLASILTLLWVRINPFVAKGGPVLEICGLNCGN"
BASE COUNT     1056 a    721 c    912 g   1139 t
ORIGIN
```

Query Match 36.3%; Score 1429.4; DB 6; Length 3828;
Best Local Similarity 65.8%; Pred. No. 4.8e-266;
Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;

```
Qy        124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
Db        219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy        184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
Db        279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy        244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
Db        339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy        304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
Db        399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy        364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
Db        459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy        424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
Db        519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy        484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
Db        579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638

Qy        544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
Db        639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT--------------CAGATTC 683

Qy        604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
Db        684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743

Qy        664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
Db        744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803

Qy        724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
Db        804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863

Qy        784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
Db        864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923

Qy        844 CGGGGAATGATGGTGGTGGTGATGATGGT--------------------------GACG 876
```


Db 924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983
Qy 877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
Db 984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043
Qy 937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
Db 1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103
Qy 997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
Db 1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163
Qy 1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
Db 1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223
Qy 1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
Db 1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283
Qy 1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
Db 1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343
Qy 1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
Db 1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403
Qy 1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
Db 1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463
Qy 1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
Db 1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523
Qy 1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
Db 1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583
Qy 1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
Db 1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643
Qy 1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
Db 1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703
Qy 1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
Db 1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763
Qy 1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db 1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823
Qy 1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db 1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883
Qy 1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
Db 1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943
Qy 1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db 1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003
Qy 1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db 2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063
Qy 1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db 2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123
Qy 2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db 2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183
Qy 2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db 2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243
Qy 2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db 2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303
Qy 2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db 2304 AGACTTCA----------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345
Qy 2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db 2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405
Qy 2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db 2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465
Qy 2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db 2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525
Qy 2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db 2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585
Qy 2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db 2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645
Qy 2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db 2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705
Qy 2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db 2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765
Qy 2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
Db 2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825
Qy 2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796
Db 2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885
Qy 2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
Db 2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945
Qy 2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
Db 2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005
Qy 2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
Db 3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065
Qy 2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
Db 3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

```
Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db       3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db       3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db       3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db       3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db       3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397
              |
Db       3483 G 3483
```

```
RESULT 11
BD022677
LOCUS       BD022677               3828 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Manipulation of cellulose and/or beta-1,4-glucan.
ACCESSION   BD022677
VERSION     BD022677.1  GI:22563900
KEYWORDS    JP 2001510326-A/5.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 3828)
  AUTHORS   Ariori,A., Williamson,R.E., Bettsunar,A.S. and Penn,L.
  TITLE     Manipulation of cellulose and/or beta-1,4-glucan
  JOURNAL   Patent: JP 2001510326-A 5 31-JUL-2001;
            THE AUSTRALIAN NATIONAL UNIVERSITY, COMMONWEALTH SCIENTIFIC AND
            INDUSTRIAL RESEARCH ORGANIZATION
COMMENT     OS   Arabidopsis thaliana (thale cress)
            PN   JP 2001510326-A/5
            PD   31-JUL-2001
            PF   24-JUN-1997 JP 1998503665
            PR   27-JUN-1996 AU      PO 0699
            PI   ANTONIO ARIORI,RICHARD EDWARD WILLIAMSON,ANDREAS STEPHAN  PI
            BETTSUNAR,
            PI   LYONKAI PENN
            PC   C12N15/09,A01H5/00,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/
            PC   10,C12N5/10,
            PC   C12N9/10,C12N15/00,C12N5/00,C12N5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             239. .3490.
FEATURES             Location/Qualifiers
     source          1. .3828
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
BASE COUNT     1056 a    721 c    912 g   1139 t
ORIGIN

  Query Match             36.3%;  Score 1429.4;  DB 6;  Length 3828;
  Best Local Similarity   65.8%;  Pred. No. 4.8e-266;
```

```
  Matches 2172;  Conservative    0;  Mismatches 1066;  Indels   63;  Gaps    4;

Qy        124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
Db        219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy        184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
Db        279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy        244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
Db        339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy        304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
Db        399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy        364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
Db        459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy        424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
Db        519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy        484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
Db        579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638

Qy        544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
Db        639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT---------------CAGATTC 683

Qy        604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
Db        684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743

Qy        664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
Db        744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803

Qy        724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
Db        804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863

Qy        784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
Db        864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923

Qy        844 CGGGGAATGATGGTGGTGGTGATGATGGT---------------------------GACG 876
Db        924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983

Qy        877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
Db        984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043

Qy        937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
Db       1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103

Qy        997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
Db       1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163

Qy       1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
Db       1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223

Qy       1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
Db       1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283
```

```
Qy      1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
Db      1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343

Qy      1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
Db      1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403

Qy      1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
Db      1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463

Qy      1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
Db      1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523

Qy      1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
Db      1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583

Qy      1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
Db      1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643

Qy      1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
Db      1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703

Qy      1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
Db      1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763

Qy      1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db      1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823

Qy      1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db      1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883

Qy      1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
Db      1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943

Qy      1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db      1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003

Qy      1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db      2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063

Qy      1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db      2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123

Qy      2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db      2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183

Qy      2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db      2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243

Qy      2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db      2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303

Qy      2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db      2304 AGACTTCA----------------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345
```

```
Qy      2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db      2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405

Qy      2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db      2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465

Qy      2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db      2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525

Qy      2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db      2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585

Qy      2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db      2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645

Qy      2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db      2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705

Qy      2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db      2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765

Qy      2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
Db      2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825

Qy      2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796
Db      2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885

Qy      2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
Db      2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945

Qy      2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
Db      2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005

Qy      2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
Db      3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065

Qy      2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
Db      3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

Qy      3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db      3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy      3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db      3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy      3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db      3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy      3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db      3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy      3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db      3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy      3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
```

```
Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397

Db       3483 G 3483
```

RESULT 12
AF200526
LOCUS AF200526 3725 bp mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-2 (CesA-2) mRNA, complete cds.
ACCESSION AF200526
VERSION AF200526.1 GI:9622875
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3725)
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S., Xoconostle-Cazares,B. and Delmer,D.P.
TITLE A comparative analysis of the plant cellulose synthase (CesA) gene family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3725)
AUTHORS Dhugga,K.S. and Helentjaris,T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES Location/Qualifiers
source 1. .3725
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"
gene 1. .3725
/gene="CesA-2"
CDS 179. .3403
/gene="CesA-2"
/note="CESA-2"
/codon_start=1
/product="cellulose synthase-2"
/protein_id="AAF89962.1"
/db_xref="GI:9622876"
/translation="MAANKGMVAGSHNRNEFVMIRHDGDAPVPAKPTKSANGQVCQIC
GDTVGVSATGDVFVACNECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGD
DEEEDVDDLDNEFNYKQGNGKGPEWQLQGDDADLSSSARHDPHHRIPRLTSGQQISGE
IPDASPDRHSIRSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRVKQD
KNMLQVTNKYPEARGDMEGTGSNGEDMQMVDDARLPLSRIVPISSNQLNLYRIVIILR
LIILCFFFQYRISHPVRNAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDRLA
LRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVSCYVSDDGS
AMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAM
KREYEEFKIRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDT
DGNELPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKA
LREAMCFMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVY
VGTGCCFNRQALYGYDPVLTEADLEPNIVVKSCCGRRKRKNKSYMDSQSRIMKRTESS
APIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPAS
LLKEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFK
GSAPINLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITS
VPLIAYCVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWW
RNEQFWVIGGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLL
IPPTTVLVINLVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQN
RTPTIVIVWSILLASIFSLLWVKIDPFISPTQKAAALGQCGVNC"
BASE COUNT 927 a 807 c 950 g 1041 t
ORIGIN

Query Match 35.0%; Score 1378.4; DB 8; Length 3725;

Best Local Similarity 66.4%; Pred. No. 3.6e-256;
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;

```
Qy        119 CGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCA 178
              ||| |   || |  ||   | | ||||| ||| | |   ||| ||||||| ||||| ||
Db        154 CGGTGTGTGTAGGAGGAGCGCGGAGATGGCGGCCAACAAGGGGATGGTGGCAGGCTCTCA 213

Qy        179 CAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA------TCCCGGGCCGAAGCC 232
              |||||||||||||| ||||| | ||||||| |||||||||      |  |  | | ||||| 
Db        214 CAACCGCAACGAGTTCGTCATGATCCGCCACGACGGCGACGCGCCTGTCCCGGCTAAGCC 273

Qy        233 GCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCC 292
              ||  |    ||| |||||||| |||||||||| ||||||   || ||| || |  |
Db        274 CACGAAGAGTGCGAATGGGCAGGTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTCAGC 333

Qy        293 CGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTA 352
              |   ||  ||   ||| || || ||||| ||||| |||||||| ||||||||     |||||
Db        334 CACTGGTGATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCTTGCTA 393

Qy        353 CGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCG 412
              || |||||||||  ||| || |  ||   |||||| |||||||||||| |||||||| |
Db        394 TGAGTACGAGCGCAAGGAAGGGAACCAATGCTGCCCTCAGTGCAAGACTAGATACAAGAG 453

Qy        413 CCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGA 472
              |  || ||  |||  || ||     ||||| || |||||||| |  || ||:||||||||||
Db        454 ACAGAAAGGTAGCCCTCGAGTTCATGGTGATGATGAGGAGGAAGATGTTGATGACCTGGA 513

Qy        473 CAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGG 532
              ||| || |||||||  |                    |   ||      | ||
Db        514 CAATGAATTCAACTATAA----------------------GCAAGGCAATGGGAAGGG 549

Qy        533 CCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAA 592
              ||    | |  |    | |||| |||| || ||    |  ||   |||  |||   |
Db        550 CCCAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCC 609

Qy        593 CCCCAAT---GTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCA 649
              | | ||     |||||| ||| || |  || |||  | |    ||          || ||
Db        610 ACACCATCGGATTCCACGCCTTACAAGTGGACAACAGATATCTGGAGAGATCCCTGATGC 669

Qy        650 GCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTA 709
              | |      | |||| ||    ||                    |   |       ||
Db        670 ATCCCCTGACCGTCATTCTATCCGCAGT------------------CCAACATCGAGCTA 711

Qy        710 TGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGC 769
              ||  ||||| ||| | || || |     |||  | |||||||  || |||||  |   | |
Db        712 TGTTGATCCAAGCGTTCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTC 771

Qy        770 ATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA 829
              ||||||  |  |||||||| |||||||| |  | ||||  |||| |   | ||| ||
Db        772 CTATGGGCTTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831

Qy        830 GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG---------------- 873
              | |        |  | || | ||     |  || | | | |
Db        832 CAAAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGGAC 891

Qy        874 --------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAA 925
                      | |  | ||| | | | |  | ||||| ||| | | ||   ||  | | |
Db        892 TGGCTCAAATGGAGAAGATATGCAAATGGTTGATGATGCACGCCTACCTTTGAGCCGCAT 951

Qy        926 AATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGT 985
              | ||| || | |||| |||| | || |  ||  ||||  | |||||| | || ||| |
Db        952 TGTGCCAATTTCCTCAAACCAGCTCAACCTTTACCGGATAGTAATCATTCTCCGTCTTAT 1011

Qy        986 GGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTT 1045
              |  || | |||||||||| || ||  | |  ||||| |||  | |||| | ||  ||
Db       1012 CATCCTGTGCTTCTTCTTCCAATATCGTATCAGTCATCCAGTGCGTAATGCTTATGGATT 1071

Qy       1046 GTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATT 1105
              ||||||  |||||||||||||||  |||||||||| |||| ||| |||| ||||| ||
Db       1072 GTGGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGCCTTGTCCTGGCTTCTAGATCAGTT 1131

Qy       1106 CCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGA 1165
              |||||| ||||  || ||  |  | ||||| || || ||| ||||  || |||||||  ||
```

```
Db       1132 CCCAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTGAGGTATGA 1191
Qy       1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225
Db       1192 TAGAGAGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTTGTCAGTACAGTGGATCC 1251
Qy       1226 CTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTA 1285
Db       1252 ATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTGCTGTGGATTA 1311
Qy       1286 TCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGA 1345
Db       1312 CCCTGTTGACAAAGTGTCATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGA 1371
Qy       1346 AGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAA 1405
Db       1372 GTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGAAGCACAA 1431
Qy       1406 TATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAA 1465
Db       1432 TATTGAACCAAGAGCTCCAGAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAA 1491
Qy       1466 GGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAA 1525
Db       1492 AATTCAACCTTCATTTGTTAAGGAAAGACGAGCAATGAAGAGAGAGTATGAAGAATTCAA 1551
Qy       1526 GGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAAT 1585
Db       1552 AATAAGAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCAT 1611
Qy       1586 GCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGT 1645
Db       1612 GGCTGATGGAACTGCTTGGCCTGGGAATAACCCTAGGGACCATCCTGGCATGATTCAGGT 1671
Qy       1646 CTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTA 1705
Db       1672 GTTCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAAATGAATTACCACGTCTTGTCTA 1731
Qy       1706 TGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGC 1765
Db       1732 TGTCTCTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAATGAATGC 1791
Qy       1766 ATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGA 1825
Db       1792 ACTGATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGTGA 1851
Qy       1826 TCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTT 1885
Db       1852 CCATTACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCAGC 1911
Qy       1886 ACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCA 1945
Db       1912 TCTAGGAAGGAAAACTTGTTATGTACAATTTCCACAAAGATTTGATGGCATTGACTTGCA 1971
Qy       1946 TGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGG 2005
Db       1972 CGATCGATATGCTAATAGGAACATAGTCTTCTTTGATATCAACATGAAAGGTCTAGATGG 2031
Qy       2006 TATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGG 2065
Db       2032 CATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGG 2091
Qy       2066 TTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTG 2125
Db       2092 ATATGAT----------------CCTGTTTTGACTGAAGCTGATCTGGAACCTAACATT 2134
Qy       2126 GTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAAC 2185
Db       2135 GTTGTTAAG----AGCTGCTGTGGTAGAAGGAAGAGAAAGAACAAGAGTTATATGGATAG 2190
Qy       2186 AGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCT 2245
Db       2191 TCAAAGCCGTA---------TTATGAAGAGAACAGAATCTTCAGCTCCCATCTTTAACAT 2241

Qy       2246 TGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCA 2305
Db       2242 GGAAGACATCGAGGAGGGTATTGAAGGTTATGAGGATGAAAGGTCAGTGCTTATGTCCCA 2301
Qy       2306 ACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGA 2365
Db       2302 GAGGAAATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTCATTGCATCCACCTTTATGAC 2361
Qy       2366 GAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGT 2425
Db       2362 TCAAGGTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTGAAGGAAGCTATCCATGT 2421
Qy       2426 CATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGG 2485
Db       2422 TATCAGCTGTGGGTACGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGATCTATGG 2481
Qy       2486 ATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAAT 2545
Db       2482 TTCAGTTACAGAGGATATTCTGACTGGGTTTAAAATGCATGCAAGAGGCTGGCAATCAAT 2541
Qy       2546 TTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCG 2605
Db       2542 CTACTGCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATCG 2601
Qy       2606 TCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTG 2665
Db       2602 TCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGCAGACATTG 2661
Qy       2666 CCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAA 2725
Db       2662 TCCTATATGGTATGGCTACAATGGGCGATTGAAGCTTTTGGAGAGGCTGGCTTACATTAA 2721
Qy       2726 CTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCAT 2785
Db       2722 CACCATTGTTTATCCAATCACATCTGTTCCGCTTATCGCCTATTGTGTGCTTCCTGCTAT 2781
Qy       2786 CTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTT 2845
Db       2782 CTGTCTTCTTACCAATAAATTTATCATTCCTGAGATTAGTAATTATGCTGGAATGTTCTT 2841
Qy       2846 CATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGT 2905
Db       2842 CATTCTTCTTTTTGCCTCCATTTTCGCAACTGGTATATTGGAGCTCAGATGGAGTGGTGT 2901
Qy       2906 TGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACA 2965
Db       2902 TGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCTCTGCCCA 2961
Qy       2966 CCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCAC 3025
Db       2962 TCTCTTCGCGGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCAACTTCAC 3021
Qy       3026 CGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATG 3082
Db       3022 AGTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTG 3081
Qy       3083 GACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGC 3142
Db       3082 GACCAGTTTGCTCATCCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAATGGTGGC 3141
Qy       3143 TGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCT 3202
Db       3142 AGGAATTTCGTATGCCATTAACAGCGGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCT 3201
Qy       3203 ATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAG 3262
Db       3202 GTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCCTCAAGGGTCTCATGGGCAG 3261
Qy       3263 GCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTC 3322
Db       3262 GCAGAACCGCACGCCAACAATCGTCATCGTTTGGTCCATCCTCCTTGCGTCTATCTTCTC 3321
```

```
Qy       3323 GCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352
Db       3322 CTTGCTGTGGGTGAAGATCGATCCTTTCAT 3351
```

RESULT 13
AF200525
LOCUS AF200525 3752 bp mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-1 (CesA-1) mRNA, complete cds.
ACCESSION AF200525
VERSION AF200525.1 GI:9622873
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3752)
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S., Xoconostle-Cazares,B. and Delmer,D.P.
TITLE A comparative analysis of the plant cellulose synthase (CesA) gene family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3752)
AUTHORS Dhugga,K.S. and Helentjaris,T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES Location/Qualifiers
source 1. .3752
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"
gene 1. .3752
/gene="CesA-1"
CDS 183. .3410
/gene="CesA-1"
/note="CESA-1"
/codon_start=1
/product="cellulose synthase-1"
/protein_id="AAF89961.1"
/db_xref="GI:9622874"
/translation="MAANKGMVAGSHNRNEFVMIRHDGDVPGSAKPTKSANGQVCQIC
GDSVGVSATGDVFVACNECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGD
EDEEDVDDLDNEFNYKQGSGKGPEWQLQGDDADLSSSARHEPHHRIPRLTSGQQISGE
IPDASPDRHSIRSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRVKQD
KNMMQVTNKYPEARGGDMEGTGSNGEXMQMVDDARLPLSRIVPISSNQLNLYRVVIIL
RLIILCFFFQYRVSHPVRDAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDRL
ALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILSVDYPVDKVSCYVSDDG
SAMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRA
MKREYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLD
TDGNELPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSK
ALREAMCFMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPV
YVGTGCCFNRQALYGYDPVLTEADLEPNIVIKSCCGRRKKKNKSYMDSQSRIMKRTES
SAPIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPA
SLLKEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCF
KGSAPINLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPIT
SIPLIAYCVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDW
WRNEQFWVIGGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSL
LIPPTTVLVINLVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQ
NRTPTIVIVWSILLASIFSLLWVKIDPFISPTQKAAALGQCGVNC"
BASE COUNT 944 a 804 c 965 g 1038 t 1 others
ORIGIN

Query Match 34.7%; Score 1364; DB 8; Length 3752;
Best Local Similarity 66.1%; Pred. No. 2.2e-253;
Matches 2165; Conservative 0; Mismatches 1001; Indels 111; Gaps 9;

```
Qy       115 TCGGCGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCT 174
Db       154 TCGGTGGGAGCAGCGGGGAGGGAGCGGAGATGGCGGCCAACAAGGGGATGGTGGCGGGCT 213
Qy       175 CCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGAT------CCCGGGCCGA 228
Db       214 CGCACAACCGCAACGAGTTCGTCATGATCCGCCACGACGGCGATGTGCCGGGCTCGGCTA 273
Qy       229 AGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTG 288
Db       274 AGCCCACAAAGAGTGCGAATGGACAGGTCTGCCAGATTTGCGGTGACTCTGTGGGTGTTT 333
Qy       289 CCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACT 348
Db       334 CAGCCACTGGTGATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCAT 393
Qy       349 GCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACA 408
Db       394 GCTATGAGTATGAGCGCAAGGAGGGGAACCAATGCTGCCCCCAGTGCAAGACTAGATACA 453
Qy       409 AGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACC 468
Db       454 AGAGACAGAAAGGTAGCCCTCGAGTTCATGGTGATGAGGATGAGGAAGATGTTGATGACC 513
Qy       469 TGGACAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCT 528
Db       514 TAGACAATGAATTCAACTACAA----------------------GCAAGGCAGTGGGA 549
Qy       529 ACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCC--- 585
Db       550 AAGGCCCAGAGTGGCAACTGCAAGGAGATGATGCTGATCTGTCTTCATCTGCTCGCCATG 609
Qy       586 AGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGG 645
Db       610 AGCCACATCATCGGATTCCACGCCTGACAAGCGGTCAACAGATATCTGGAGAGATTCCTG 669
Qy       646 AGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTC 705
Db       670 ATGCTTCCCCTGACCGTCATTCTATCCGCAGT----------------CCAACATCGA 711
Qy       706 CTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTG 765
Db       712 GCTATGTTGATCCAAGCGTCCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGA 771
Qy       766 CTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGAC 825
Db       772 ATTCCTATGGGCTTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAAC 831
Qy       826 AAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG------------ 873
Db       832 AGGACAAAAATATGATGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGGAGACATGG 891
Qy       874 ---------------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGT 918
Db       892 AGGGGACTGGCTCAAATGGAGAANATATGCAAATGGTTGATGATGCACGGCTACCTTTGA 951
Qy       919 CCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTC 978
Db       952 GCCGTATCGTGCCAATTTCCTCAAACCAGCTCAACCTTTACCGGGTAGTGATCATTCTCC 1011
Qy       979 GGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCAT 1038
Db      1012 GTCTTATCATCCTGTGCTTCTTCTTCCAGTATCGTGTCAGTCATCCAGTGCGTGATGCTT 1071
Qy      1039 TTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTG 1098
Db      1072 ATGGATTATGGCTAGTATCTGTTATCTGCGAGGTCTGGTTTGCCTTGTCTTGGCTTCTAG 1131
Qy      1099 ATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGA 1158
Db      1132 ATCAGTTCCCAAAATGGTATCCAATCAACCGTGAGACATATCTTGACAGGCTTGCATTGA 1191
Qy      1159 GGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGG 1218
```

```
Db      1192 GGTATGATAGAGAGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTCGTCAGTACAG 1251

Qy      1219 TTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGG 1278

Db      1252 TGGATCCATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTTCTG 1311

Qy      1279 TGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAA 1338

Db      1312 TGGATTACCCTGTTGACAAAGTGTCATGCTATGTTTCTGATGATGGTTCAGCTATGCTGA 1371

Qy      1339 CGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAAC 1398

Db      1372 CTTTTGAGTCTCTCTCAGAAACCGCAGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGA 1431

Qy      1399 GGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGA 1458

Db      1432 AGCACAATATTGAACCAAGAGCTCCAGAATTTTACTTTGCTCAAAAAATAGATTACCTGA 1491

Qy      1459 AAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGG 1518

Db      1492 AGGACAAAATTCAACCTTCATTTGTTAAGGAAAGACGCGCAATGAAGAGGGAGTATGAAG 1551

Qy      1519 AATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGAT 1578

Db      1552 AATTCAAAGTAAGAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGT 1611

Qy      1579 GGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGA 1638

Db      1612 GGACCATGGCTGATGGAACTGCATGGCCTGGGAATAATCCTAGGGACCATCCTGGCATGA 1671

Qy      1639 TTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGAT 1698

Db      1672 TTCAGGTTTTCTTGGGGCACAGTGGTGGGCTCGACACTGATGGAAATGAGTTACCACGTC 1731

Qy      1699 TGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTA 1758

Db      1732 TTGTCTATGTCTCTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAA 1791

Qy      1759 TGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGG 1818

Db      1792 TGAATGCGCTGATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGG 1851

Qy      1819 ATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGG 1878

Db      1852 ATTGCGACCATTACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGG 1911

Qy      1879 ACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTG 1938

Db      1912 ATCCGGCTCTAGGAAGGAAAACTTGTTATGTACAATTTCCACAGAGATTTGATGGCATTG 1971

Qy      1939 ATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTT 1998

Db      1972 ACTTGCACGATCGATATGCTAATCGGAACATAGTTTTCTTTGATATCAACATGAAAGGTC 2031

Qy      1999 TGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCAT 2058

Db      2032 TGGATGGCATTCAGGGTCCAGTTTACGTGGGAACAGGATGCTGTTTCAATAGACAGGCTT 2091

Qy      2059 TATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGC 2118

Db      2092 TGTATGGATACGAT-----------------CCTGTTTTGACTGAAGCTGATCTGGAGCC 2134

Qy      2119 CCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAAC 2178

Db      2135 AAACATTGTTATTAAG----AGCTGCTGTGGTAGAAGGAAGAAAAAGAACAAGAGTTATA 2190

Qy      2179 CCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCAT 2238

Db      2191 TGGATAGTCAAAGCCGTA---------TTATGAAGAGAACAGAATCTTCAGCTCCCATCT 2241

Qy      2239 ATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTG 2298
```

```
Db      2242 TCAATATGGAAGACATCGAAGAGGGTATTGAAGGTTACGAGGATGAAAGGTCAGTGCTTA 2301

Qy      2299 TAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACAC 2358

Db      2302 TGTCCCAGAGGAAATTGGAGAAACGCTTTGGTCAGTCTCCTATTTTCATTGCATCCACCT 2361

Qy      2359 TTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTA 2418

Db      2362 TTATGACACAAGGTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTAAAGGAAGCTA 2421

Qy      2419 TACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGA 2478

Db      2422 TCCATGTCATCAGTTGTGGATATGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGA 2481

Qy      2479 TCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGC 2538

Db      2482 TCTATGGTTCAGTAACGGAGGATATTCTGACTGGGTTTAAAATGCATGCAAGGGGCTGGC 2541

Qy      2539 GGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTT 2598

Db      2542 AATCAATCTACTGCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTT 2601

Qy      2599 CAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCA 2658

Db      2602 CCGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGTA 2661

Qy      2659 ATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCT 2718

Db      2662 GACATTGTCCTATCTGGTATGGTTACAATGGACGATTGAAGCTTTTGGAGAGGCTGGCTT 2721

Qy      2719 ACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGC 2778

Db      2722 ACATCAACACTATTGTATATCCAATCACATCCATTCCGCTTATTGCCTATTGTGTGCTTC 2781

Qy      2779 CTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCC 2838

Db      2782 CCGCTATCTGCCTCCTTACCAATAAATTTATCATTCCTGAGATTAGCAATTATGCTGGGA 2841

Qy      2839 TGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGA 2898

Db      2842 TGTTCTTCATTCTTCTTTTCGCCTCCATTTTTGCCACTGGTATATTGGAGCTTAGATGGA 2901

Qy      2899 GTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGT 2958

Db      2902 GTGGTGTTGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCT 2961

Qy      2959 CCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAA 3018

Db      2962 CTGCCCATCTCTTCGCAGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCA 3021

Qy      3019 GCTTCACCGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACAT 3075

Db      3022 ACTTCACAGTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGT 3081

Qy      3076 TCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTG 3135

Db      3082 TCAAGTGGACCAGTTTGCTCATTCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAA 3141

Qy      3136 TGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTG 3195

Db      3142 TGGTGGCAGGAATTTCTTATGCCATTAACAGTGGCTACCAATCCTGGGGTCCGCTCTTTG 3201

Qy      3196 GGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGG 3255

Db      3202 GAAAGCTGTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCCTCAAGGGTCTCA 3261

Qy      3256 TTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAA 3315

Db      3262 TGGGAAGGCAGAACCGCACACCAACAATCGTCATTGTCTGGTCCATCCTTCTTGCATCTA 3321

Qy      3316 TCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352

Db      3322 TCTTCTCCTTGCTGTGGGTGAAGATCGATCCTTTCAT 3358
```

```
RESULT 14
AC104487/c
LOCUS       AC104487               171376 bp    DNA     linear   PLN 15-FEB-2003
DEFINITION  Oryza sativa chromosome 3 BAC OSJNBa0042I09 genomic sequence,
            complete sequence.
ACCESSION   AC104487
VERSION     AC104487.3  GI:28269486
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 171376)
  AUTHORS   Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Gansberger,K., Jones,K.M.,
            Overton II,L.L., Tsitrin,T., Kim,M.M., Bera,J.J., Jin,S.S.,
            Fadrosh,D.W., Tallon,L.J., Koo,H., Zismann,V., Hsiao,J., Blunt,S.,
            Vanaken,S.S., Riedmuller,S.B., Utterback,T.T., Feldblyum,T.V.,
            Yang,Q.Q., Haas,B.J., Suh,B.B., Peterson,J.J., Quackenbush,J.,
            White,O., Salzberg,S.L. and Fraser,C.M.
  TITLE     Oryza sativa chromosome 3 BAC OSJNBa0042I09 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171376)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-DEC-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 171376)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   4  (bases 1 to 171376)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
REFERENCE   5  (bases 1 to 171376)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT     On Feb 7, 2003 this sequence version replaced gi:18497084.
            Address all correspondence to:rice@tigr.org

            BAC clone OSJNBa0042I09 is from Oryza sativa chromosome 3
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.
            Genes were identified by a combination of several methods: Gene
            prediction programs including Fgenesh (http://www.softberry.com/),
            genscan and Genscan+ (Chris Burge,
            http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM (Mark Borodovsky,
            http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer
            (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
            searches of the complete sequence against a peptide database and
            the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).
            Annotated genes are named to indicate the level of evidence for
            their annotation. Genes with similarity to other proteins are named
            after the database hits. Genes without significant peptide
            similarity but with EST similarity are named as  unknown  proteins.
            Genes without protein or EST similarity, that are predicted by more
            than two gene prediction programs over most of their length are
            annotated as  hypothetical  proteins. Genes encoding tRNAs are
            predicted by tRNAscan-SE (Sean Eddy,
            http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
            identified by repeatmasker (Arian Smit,
            http://ftp.genome.washington.edu/RM/RepeatMasker.html). This BAC
            overlaps with rice BAC OSJNBa0096I06 (AC092557) and OSJNBa0075M12
            (AC096856).
FEATURES             Location/Qualifiers
     source          1. .171376
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="3"
                     /map="S11301"
                     /clone="OSJNBa0042I09"
                     /note="japonica cultivar-group"
     repeat_region   387. .407
                     /rpt_family="AT_rich"
     repeat_region   1475. .1511
                     /rpt_family="AT_rich"
     gene            2280. .6614
                     /gene="OSJNBa0042I09.1"
                     /note="similar to Pto kinase interactor 1 GB:AAC61805
                     GI:3668069 (Lycopersicon esculentum); EST C72353,BF475198,
                     BF475196, BE039861 from this gene"
     mRNA            join(2280. .2606,3008. .3129,3242. .4527,4612. .4860,
                     4958. .5124,5321. .5400,5486. .5633,5858. .6007,6091. .6614)
                     /gene="OSJNBa0042I09.1"
     CDS             join(2280. .2606,3008. .3129,3242. .4527,4612. .4860,
                     4958. .5124,5321. .5400,5486. .5633,5858. .6007,6091. .6327)
                     /gene="OSJNBa0042I09.1"
                     /codon_start=1
                     /product="putative serine/threonine kinase"
                     /protein_id="AAO41118.1"
                     /db_xref="GI:28376688"
                     /translation="MEEGAMDGAGGGVAEEGCSGEEREGVADVAVGGGGGEGEREGRT
                     LVVGVRADAASRTLLTWTFINVAAPGDRIVAVHVVLASAPVAAATTAVDFDTMLAVYE
                     GFCNLKQINLKLKICKDSSVRKALVREANLFGASKVIVGIAKKKRGISSLHSVAKYCS
                     KKLPAKCAVLAVDSGKIVFRRESNVHSGKVSAEIPGCGDNELYCEVPFLARQCKEEPL
                     PLHEPPRDGGGSAGEEEEEHDVGTKETEPVNAVSGEQQPSGVDPAELSPDQVQGDVDP
                     SDKGEESTADQKDEISELPGEGASVLYCVLPERNGHSAASTSSRQSNDSTEPPTEGNG
                     ELYCILPPRNDHSGRSSGDSSRSTASRKHDDSASLSAEGDGELYCRLPRTGHSGRSSG
                     GSKRSFGAKGLIRRSSSFSSDIQKDVSVCTTTTEQTSSMVSTEAEDSPKNAARDVDTP
                     SSSPMSLRRMIEGRPDRCRLRRRIFNHQRSSSFEWAKISMVQWAMRLPSRYTSVADNK
                     SFKSDASPRLNCDSECESTSAVDTESMFSFSLYDISWPPNEVESLQEKYSSTCRLFTY
                     EELKLATSNFSPDMLIGKGGTSQVYKAQLFDGTLSAVKILKPSVDAIQEFVTEVEIAT
                     SLQHDNIVSLRGFSSDNYSLVLVYDYMLQGSLDKALHGKHDSKDSLSWEKRNKIAIGI
                     AKALEYLHHGSVTQSVIHGDVKSSNILLSEDFQAQLCDFGLAKQVSASTPHLTCTDIT
                     GTFGYLAPEYFSHGKVNEKIDVYAFGVVILEIISGRRPIRTGCSKGQESLVGWAKPLL
                     SSGEIKQLVDPFLGNDYDCDEMERMTLAASLCTRTSSHSRPEMSQMLKLLEGDDETIH
                     WARTQVTASFDGSDEEAVAAPDSNMQSHLNLALLGVEDDTLSHCSTEQTMDTSADGYW
                     SRSSSFD"
     repeat_region   2475. .2556
                     /rpt_family="(CCG)n"
     gene            8631. .9257
                     /gene="OSJNBa0042I09.2"
                     /note="EST CA756434 from this gene"
     mRNA            8631. .9257
                     /gene="OSJNBa0042I09.2"
     CDS             8636. .9058
                     /gene="OSJNBa0042I09.2"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAO41120.1"
                     /db_xref="GI:28376690"
                     /translation="MALDESFKRPGTIPFKWELQPGVPKQQPHHGGGGAAVGSSSASS
                     SPAPAPAAASTLLLPPRLLAPPPAAAASHGGCDTGANILASTTTPSPSSSSHRRSMSA
                     RFTASLVLPFTRPRRGRSANSKDEDDIAFTVLYGDKIV"
     repeat_region   9129. .9159
                     /rpt_family="AT_rich"
     gene            9981. .11794
                     /gene="OSJNBa0042I09.3"
                     /note="predicted by fgenesh"
     mRNA            join(<9981. .10658,10805. .10951,11258. .11360,11505. .>11794)
                     /gene="OSJNBa0042I09.3"
     CDS             join(9981. .10658,10805. .10951,11258. .11360,11505. .11794)
                     /gene="OSJNBa0042I09.3"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAO41131.1"
                     /db_xref="GI:28376701"
```

```
                /translation="MFFHLHARARRRKNFITRLQDGDQLAVSHDELQELASSHFGAIL
                GTSVERDFALDLHSLGLCPAHLGGLDNPFSEEEVWAVVKALPSDKAPGPDGFTGRFYK
                TCWSVIKPEVMAALNSLFLANGRGFSSLNDTLISLLPKKDEAVDVKDFRPISPIHSFG
                KLFSKIPASRLSPHLGELVAPNQSASSKADPFHDNFRYVQLAARALHAKHIPRLLLKV
                DIAKAFDTGDPLSPMLFILVMDVLNTMLRKASESGGFLPLNDHASLYADDLVLFLSPS
                VLSSIPVHVSMVIGLPAWMIKAINKKRRAFLWTGNPYWDGLKAPAERAVTDMFQASTF
                FVPGDGESILFWSDQWIGGRSVASLAPDLLFVVPPRLRGSRTVSSGLANNAWISDIRG
                ALTVPGISQFLLV"
gene            12826. .15886
                /gene="OSJNBa0042I09.4"
                /note="EST BF475198, BF475196, BE039861 from this gene"
mRNA            join(<12826. .12909,13348. .13725,14212. .14361,14548. .14627,
                14713. .14866,15074. .15165,15211. .15260,15347. .15886)
                /gene="OSJNBa0042I09.4"
CDS             join(12826. .12909,13348. .13725,14212. .14361,14548. .14627,
                14713. .14866,15074. .15165,15211. .15260,15347. .15363)
                /gene="OSJNBa0042I09.4"
                /codon_start=1
                /product="unknown protein"
                /protein_id="AAO41135.1"
                /db_xref="GI:28376705"
                /translation="MQLNPKARTRRNKTKVKRVAMRRDCIERDRLVWRWTSDQRYSAH
                SAYQAFFLGQHSFACVDLLLNAKGPAKCKFFLWFAFQQRCWMADLLQKRSIDSHSTCP
                FCAQELETANHILLDYVFARQVWLRVLSPPGWNALSPPVAAGSRTGGHLPRLAAPVAG
                EVFKIAIGIAKALEYLHNGSITQSVIHGDVKSSNILLSEDFQAQLCDFGLAKQVSAST
                PRLTCTDITGTFGYLAPEYFSHGKVNEKIDVYAFGVVILVLEIISGRRHIRTGCAKGQ
                ESLVGWAKPLLSSGEIKQLVDPFLGNDYDCDEMERMILAMHKDIFSFPTGNVTDFETA
                "
repeat_region   16877. .16942
                /rpt_family="tandem"
repeat_region   16879. .16944
                /rpt_family="CT-rich"
repeat_region   17388. .17570
                /rpt_family="tandem"
repeat_region   complement(18185. .18241)
                /rpt_family="Oryza"
repeat_region   18186. .18246
                /rpt_family="Oryza"
repeat_region   18186. .18243
                /rpt_family="Oryza"
misc_feature    18186. .18241
                /note="Region  with dinucleotide (AT) repeat (28-29
                copies)"
repeat_region   18186. .18241
                /rpt_family="tandem"
repeat_region   18332. .18414
                /rpt_family="tandem"
repeat_region   18340. .18412
                /rpt_family="Oryza"
repeat_region   18343. .18414
                /rpt_family="Oryza"
repeat_region   complement(18346. .18414)
                /rpt_family="Oryza"
repeat_region   18348. .18420
                /rpt_family="Oryza"
repeat_region   complement(18348. .18415)
                /rpt_family="Oryza"
misc_feature    18348. .18413
```

Query Match 33.9%; Score 1335.4; DB 8; Length 171376;
Best Local Similarity 67.5%; Pred. No. 1e-247;
Matches 2575; Conservative 0; Mismatches 456; Indels 786; Gaps 18;

```
Qy        893 AATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAA 952
              |||||||||||| ||||| | | ||||||| || |||||  || | ||||||| | | ||
Db     139935 AATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATTCCGATTTCTTCAAGCCTGGTCAA 139876

Qy        953 TCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCG 1012
              ||| ||||||||||||||||| |||||||| |||||| |||| ||||||||||||| ||
Db     139875 TCCTTATAGGATGATTATCATAATTCGGCTGGTGGTTCTGGGCTTCTTCTTCCACTATCG 139816

Qy       1013 AGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAAT 1072
              |||||||||||| |||  |||||||||||| || |||||||||| || ||||||||||
```

```
Db     139815 AGTGATGCATCCAGTGCCTGATGCATTTGCATTATGGCTCATATCCGTCATCTGTGAAAT 139756

Qy       1073 CTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGA 1132
              |||||| |||||||||||||||||||||||| |||||||||||| |||||||||| ||
Db     139755 CTGGTTCGCCATGTCTTGGATTCTTGATCAGTTCCCAAAGTGGTTTCCTATTGAGAGGGA 139696

Qy       1133 GACTTACCTAGACCGGCTGTCACTG---------------------------------- 1157
                |||||||| |||||| || || ||
Db     139695 AACTTACCTTGACCGGTTGACATTGAGGTTTGTTACTCTTGCCACCTTGTATCACTTGCC 139636

Qy       1158 ----------------------------------------------AGGTTCGACAAGG 1170
                                                            ||||| |||||||
Db     139635 TTCTATTCTTTCCGACGTTCTGTGACTTACTATTATTTTTCCTGGATAGGTTTGACAAGG 139576

Qy       1171 AAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAA 1230
              |||||||||| |||||||||  || ||| ||||||| |||||||||||||||||||| | |
Db     139575 AAGGCCAGCAATCTCAACTCGCCCCAGTTGATTTTTTTGTCAGTACGGTTGATCCCATGA 139516

Qy       1231 AGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTG 1290
              |||| ||||| ||||||||| |||||||||||| ||||||||| |||| |||||||| |
Db     139515 AGGAGCCTCCCTTGGTCACAGCAAATACTGTTCTGTCTATCCTTGCGGTAGATTATCCGG 139456

Qy       1291 TTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCAT 1350
              | || ||||| |||||||||||||||||||||||||| ||||| ||||| || ||||||||||
Db     139455 TCGACAAGGTCTCTTGCTATGTTTCTGATGATGGGGCTGCCATGCTGACATTTGAAGCAC 139396

Qy       1351 TATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTG 1410
              | |||||||||||||||||||||||||| |||||||||||||||||| |||||| | | |
Db     139395 TGTCTGAAACATCTGAATTTGCAAAGAAGTGGGTTCCTTTCTGCAAAAGGTACAGTCTCG 139336

Qy       1411 AACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGG 1470
              | ||||| || ||||| |||||||||||||||||||||||||| ||||||| |||||||
Db     139335 AGCCTCGTGCCCCAGAATGGTACTTCCAACAGAAGATAGACTACCTGAAAGATAAGGTGG 139276

Qy       1471 CAGCAAACTTTGTTAGGGAGAGGAGAGCAATGA---------------------------- 1503
              || |||||||||| |||||||||||||||||||
Db     139275 CACCAAACTTTGTCAGGGAGAGGAGAGCAATGAAGGTACATTGCATTTCTGTTGTGTTTT 139216

Qy       1504 ------------------------------------------------------------AGA 1506
                                                                           |||
Db     139215 GCTATGATATGTAGCTCATAAAATTAATGCTAATTTTGACCTCAATTCCTTTCTCTCAGA 139156

Qy       1507 GAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTC 1566
              | || |||||||||||||||| |||||||||| || |||| ||||||||||||||||
Db     139155 GGGAATATGAGGAATTCAAGGTCAGAATCAATGCATTGGTTGCTAAAGCCCAGAAAGTTC 139096

Qy       1567 CTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATC 1626
              |||| |||||||||||||||| |||||||| |||||||||||||||| ||||||||||
Db     139095 CTGAGGAAGGATGGACAATGCAGGATGGAACTCCCTGGCCTGGAAACAACGTTCGTGATC 139036

Qy       1627 ATCCTGGAATGATTC---------------------------------------------- 1641
              |||||||||||||||
Db     139035 ATCCTGGAATGATTCAGGTATGCTTTTACAGTTCATGGTGCTGTTTAATTATCTGCTGAC 138976

Qy       1642 ------------------------------------------------------------ 1641

Db     138975 ATGATGTGATGGCTACTTTCCCCTTCTCTTGCTTTACCATGACAAATCTCAAACTTTTGC 138916

Qy       1642 ------AGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCAC 1695
                    |||||||||||||||||| || |||| |||    || |||||||||||||| |
Db     138915 AACCATAGGTCTTCCTTGGCCAAAGTGGTGGCCATGATGTGGAAGGAAATGAACTGCCTC 138856

Qy       1696 GATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTG 1755
              |||||||||||||||| ||||| ||||| |||||||||||||||||||||| |||||||
Db     138855 GATTGGTTTATGTTTCAAGAGAAAAACGGCCAGGCTATAACCATCATAAGAAGGCTGGTG 138796

Qy       1756 CTATGAATGCATT---------------------------------------------- 1768
              ||||||| |||||
Db     138795 CTATGAACGCATTGGTAAAGACAACACGTTCAAGGCATTTGTCAGTCTTCTATAGCCTGT 138736

Qy       1769 ----------------------------------------------GGTCCGAGTCTC 1780
                                                            ||||||||||||
Db     138735 AGCAGCATCCAGTGGATATTGATGTCCTTTCTGTTTGTCTGTTTTTCAGGTCCGAGTCTC 138676
```

```
Qy       1781 TGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAA 1840
Db     138675 CGCTGTACTAACTAATGCTCCATATATGTTGAACTTGGATTGTGATCACTACATCAACAA 138616

Qy       1841 CAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGT 1900
Db     138615 CAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGATCCCTTGGTAGGAAAGAAAGT 138556

Qy       1901 TTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAA 1960
Db     138555 TTGTTATGTGCAGTTCCCACAAAGATTCGATGGCATTGACCGTCATGATCGATATGCTAA 138496

Qy       1961 CCGGAATGTTGTCTTTTTTGAT---------------------------------------- 1982
Db     138495 TAGGAATGTCGTCTTTTTCGATGTAAGATTTATCTCCAATATTTCTGATTCTGTTCATCG 138436

Qy       1983 ----------------------------------------------------------ATCAAC 1988
Db     138435 TGTATGTTGAGCCATCAAACTTGATGCTGTTTCCCTATTATTTATTTCTTGCAGATCAAC 138376

Qy       1989 ATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGA 2048
Db     138375 ATGAAAGGTTTGGATGGTATTCAAGGCCCCATCTATGTCGGTACTGGATGTGTCTTTAGA 138316

Qy       2049 AGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGC 2108
Db     138315 AGACAGGCGCTGTATGGTTATGATGCCCCCAAATCAAAGAAACCACCATCAAGGACTTGC 138256

Qy       2109 AACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGG---AAGCAAAAG 2165
Db     138255 AACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGCTTTGGTAACAGGACAAACAAGAAG 138196

Qy       2166 AAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAAC 2225
Db     138195 AAGACTGCTAAACCTAAAACAGAGAAGAAAAAGAGGTTGTTTTTCAAGAGAGCAGAGAAT 138136

Qy       2226 CAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTC----------------- 2268
Db     138135 CAGTCGCCTGCATATGCTCTTGGTGAGATTGACGAAGGTGCTCCAGGTATCACTTTAACT 138076

Qy       2269 ------------------------------------------------------------ 2268
Db     138075 TCCAAGAAAACCCTGTCGTTATTCTAGATGCTTCCTCATCTTCATTGTATTAGCACACCT 138016

Qy       2269 -------------------CAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309
Db     138015 AATCTTGGAATGGTTTCTGCAGGAGCTGAAAATGAAAAGGCTGGTATTGTAAATCAACAG 137956

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369
Db     137955 AAGTTAGAAAAGAAATTTGGTCAATCTTCCGTTTTTGTTGCGTCCACACTTCTTGAGAAT 137896

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429
Db     137895 GGTGGAACCCTGAAGAGTGCAAGCCCAGCATCACTCTTAAAGGAAGCTATACATGTCATC 137836

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAG---------------------- 2466
Db     137835 AGTTGTGGTTACGAGGACAAGACAGACTGGGGAAAAGAGGTAAAATTTGGAGCATGGTTG 137776

Qy       2467 ------------------------------------------------------------ 2466
Db     137775 TGTGTTAGTTGGCAGGTCGAATATCTTCCGTTACCTTGTAAACCATAACTTCTTTGTCTT 137716

Qy       2467 -----AGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGAT 2521
Db     137715 TTTGTAGATTGGCTGGATCTACGGATCCGTTACAGAGGATATTCTTACTGGGTTTAAGAT 137656

Qy       2522 GCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTC 2581
Db     137655 GCACTGCCATGGTTGGCGGTCGATTTACTGCATACCTAAGCGGGCTGCATTCAAAGGTTC 137596
```

```
Qy       2582 TGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTAT 2641
Db     137595 TGCACCTCTCAATCTTTCTGACCGTCTTCACCAGGTTCTTCGGTGGGCGCTTGGATCTAT 137536

Qy       2642 TGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATT 2701
Db     137535 TGAAATTTTCTTCAGCAACCATTGCCCTCTTTGGTATGGCTATGGAGGTGGATTGAAATG 137476

Qy       2702 TTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTT 2761
Db     137475 TTTGGAAAGATTTTCCTACATTAATTCCATCGTGTATCCTTGGACATCCATTCCACTATT 137416

Qy       2762 GGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAG---- 2817
Db     137415 GGCTTACTGTACATTGCCTGCCATCTGTTTGCTGACAGGAAAATTTATCACTCCAGAGGT 137356

Qy       2818 ------------------------------------------------------------ 2817
Db     137355 AACGCTTATTCTGCTCGTAATTCTGTGTGGTGTGAAATTGTAGAAGACCAATGACCTGAA 137296

Qy       2818 -------------------------------AGCTGAATAATGTTGCCAGCCTGTGGTTC 2846
Db     137295 CCTTACTTTTCCTCCTTTTGTTCCATTCTGCAGCTTACTAATATCGCCAGCCTTTGGTTC 137236

Qy       2847 ATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTT 2906
Db     137235 ATGTCACTATTTATCTGCATTTTTGCTACGGGCATTCTGGAAATGAGATGGAGTGGTGTA 137176

Qy       2907 GGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACAC 2966
Db     137175 GGCATCGATGATTGGTGGAGGAACGAGCAGTTTTGGGTCATTGGAGGTGTCTCTTCACAT 137116

Qy       2967 CTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACC 3026
Db     137115 CTCTTTGCTGTGTTCCAAGGACTTCTCAAGGTCATAGCTGGTATCGATACAAGCTTCACT 137056

Qy       3027 GTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACT 3086
Db     137055 GTGACATCAAAGGGTGGAGATGATGAAGAGTTCTCAGAGCTGTATACCTTCAAATGGACA 136996

Qy       3087 ACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGC 3146
Db     136995 ACCTTGTTAATACCCCCTACCACTTTGCTCCTATTGAACTTCATCGGGGTAGTGGCTGGC 136936

Qy       3147 GTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTC 3206
Db     136935 GTTTCCAACGCAATCAATAACGGATATGAATCATGGGGGCCGCTCTTTGGGAAGCTGTTC 136876

Qy       3207 TTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAA 3266
Db     136875 TTTGCATTTTGGGTGATTGTCCATCTGTATCCGTTCTTGAAGGGTTTGGTTGGAAGGCAA 136816

Qy       3267 AACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTC 3326
Db     136815 AACAGGACACCAACAATTGTCATTGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTC 136756

Qy       3327 CTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGT 3386
Db     136755 CTCTGGGTCCGGATTGATCCTTTCCTGGCAAAGAATGACGGCCCTCTTCTGGAGGAGTGT 136696

Qy       3387 GGTTTGGATTGCAACTAGGATGTCAGTGCATCAGCTCCCCCAATC----TGCATATGCTT 3442
Db     136695 GGTTTGGATTGCAACTAGAAGGTCAGCATGTGAACTTCCCCAATCAAGTGTGTGCTTGAA 136636

Qy       3443 GAAGTATATTTTCTGGTGTTTGTCCCCATA--------------TTCAGTGTCTGTAGAT 3488
Db     136635 GAAGTATTTTTGCAGGTGTTTGTGCCCATATTTTTTTTCATTTTTTCCTCTGTAGATAG 136576

Qy       3489 AAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAATATCTGAG 3548
Db     136575 AAACACGAAAAAAAAACTTGCATTTCTTTTGTTCCATGGTG--GTAGCTCAGCCTCATAT 136518

Qy       3549 AGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGC--CGTGGAAT 3606
```

```
Db     136517 CTGCAATATTGAGCAAACGGGAGGCTGCGACAA-CCTTGTGCAGTTCGGCCACGAAAAAT 136459

Qy       3607 ACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATATAGATG 3666
              ||||   |||| ||||||||||||| ||||||||||||||| || |  | ||  | |||
Db     136458 ACAGTGAATGCCAGTGTTTGATTGTTCAGCATTCTTTATTAGTTTGGTGGAAATTTGATC 136399

Qy       3667 GGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAAC------- 3719
              || |||||||||||| || ||   |||||||||| ||||| |    |   ||
Db     136398 GGGTGAGCCGAACAG-AAAGTTGCTTGATTCTGCATTGCTCTCCACCATCTACAAAAACT 136340

Qy       3720 --TTGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCC-AGTGGAACAGAGCAACCTGCA 3776
                |  || |   | |   |||||||||| |||||||   ||||| |||| ||  ||||
Db     136339 CGGTACTTGTTTCTCAATAAGGCAGGAATGGATCTGCCGTCTGGAATAGAGGAAGATGCA 136280

Qy       3777 CATTATTTATGTA------TGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTA 3830
              |||||||||||||      |      |   || |||||||||||||| | ||||| |||
Db     136279 CATTATTTATGTAGCGAGTTAAAATATGCCTGAAGGGCTTGTTCATTCCTTGTTCCTCTT 136220

Qy       3831 TACTAGAAAAAACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCC 3890
              |    |      |  | |   ||   |||| |   ||   |  || | |  |||
Db     136219 TCTATTTCATTTTTTTGTCCTTGTTAGAACAAATAGATCTCTAGTTTTAGTGATTTAGCC 136160

Qy       3891 TGTTTTTTGTTGTGTACTGTAATCATCTGAGTTGGTT 3927
                |    ||||||    |||||   |     || |||
Db     136159 CATCCGGTGTTGTAATTTGTAAAGGTGCCCCTTTGTT 136123
```

RESULT 15
BT008654
LOCUS BT008654 3763 bp mRNA linear PLN 16-MAY-2003
DEFINITION Arabidopsis thaliana clone RAFL09-89-G08 (R19778) putative cellulose synthase catalytic subunit (RSW1) (At4g32410) mRNA, complete cds.
ACCESSION BT008654
VERSION BT008654.1 GI:30794042
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 3763)
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Full Length cDNA Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3763)
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

```
FEATURES             Location/Qualifiers
     source          1. .3763
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL09-89-G08 (R19778)"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert.
                     ecotype: Columbia"
     gene            1. .3763
                     /gene="At4g32410"
     5'UTR           1. .144
                     /gene="At4g32410"
     misc_difference 2
                     /gene="At4g32410"
                     /note="compared to genomic sequence"
                     /replace="t"
     CDS             145. .3390
                     /gene="At4g32410"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative cellulose synthase catalytic subunit
                     (RSW1)"
                     /protein_id="AAP40467.1"
                     /db_xref="GI:30794043"
                     /translation="MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMNGQICQICGD
                     DVGLAETGDVFVACNECAFPVCRPCYEYERKDGTQCCPQCKTRFRRHRGSPRVEGDED
                     EDDVDDIENEFNYAQGANKARHQRHGEEFSSSSRHESQPIPLLTHGHTVSGEIRTPDT
                     QSVRTTSGPLGPSDRNAISSPYIDPRQPVPVRIVDPSKDLNSYGLGNVDWKERVEGWK
                     LKQEKNMLQMTGKYHEGKGGEIEGTGSNGEELQMADDTRLPMSRVVPIPSSRLTPYRV
                     VIILRLIILCFFLQYRTTHPVKNAYPLWLTSVICEIWFAFSWLLDQFPKWYPINRETY
                     LDRLAIRYDRDGEPSQLVPVDVFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVACYV
                     SDDGSAMLTFESLSETAEFAKKWVPFCKKFNIEPRAPEFYFAQKIDYLKDKIQPSFVK
                     ERRAMKREYEEFKVRINALVAKAQKIPEEGWTMQDGTPWPGNNTRDHPGMIQVFLGHS
                     GGLDTDGNELPRLIYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYF
                     NNSKAIKEAMCFMMDPAIGKKCCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGI
                     QGPVYVGTGCCFNRQALYGYDPVLTEEDLEPNIIVKSCCGSRKKGKSSKKYNYEKRRG
                     INRSDSNAPLFNMEDIDEGFEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIP
                     PTTNPATLLKEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCN
                     PPRPAFKGSAPINLSDRLNQVLRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINT
                     IVYPITSIPLIAYCILPAFCLITDRFIIPEISNYASIWFILLFISIAVTGILELRWSG
                     VSIEDWWRNEQFWVIGGTSAHLFAVFQGLLKVLAGIDTNFTVTSKATDEDGDFAELYI
                     FKWTALLIPPTTVLLVNLIGIVAGVSYAVNSGYQSWGPLFGKLFFALWVIAHLYPFLK
                     GLLGRQNRTPTIVIVWSVLLASIFSLLWVRINPFVDANPNANNFNGKGGVF"
     3'UTR           3391. .3763
                     /gene="At4g32410"
     misc_difference 3747
                     /gene="At4g32410"
                     /note="not present in genomic sequence"
BASE COUNT      997 a    768 c    892 g   1106 t
ORIGIN
```

Query Match 33.9%; Score 1332.4; DB 8; Length 3763;
Best Local Similarity 65.0%; Pred. No. 2.8e-247;
Matches 2117; Conservative 0; Mismatches 1066; Indels 75; Gaps 7;

```
Qy        127 GTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCA 186
              |||     || ||  |||||||||| || |||||  |||| || || ||| ||     | |
Db        128 GTGTGTCGGTGGCTGCGATGGAGGCCAGTGCCGGCTTGGTTGCTGGATCCTACCGGAGAA 187

Qy        187 ACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGA 246
              ||||||||||     ||||| |  ||    |||  |||| | || ||   |  | |   ||
Db        188 ACGAGCTCGTTCGGATCCGACATGAATCTGATGGCGGGACCAAACCTTTGAAGAATATGA 247

Qy        247 ACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCT 306
```

```
Db       248 ATGGCCAGATATGTCAGATCTGTGGTGATGATGTTGGACTCGCTGAAACTGGAGATGTCT 307
Qy       307 TCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCC 366
Db       308 TTGTCGCGTGTAATGAATGTGCCTTCCCTGTGTGTCGGCCTTGCTATGAGTACGAGAGGA 367
Qy       367 GGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCC 426
Db       368 AAGATGGAACTCAGTGTTGCCCTCAATGCAAGACTAGATTCAGACGACACAGGGGGAGTC 427
Qy       427 AACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACT 486
Db       428 CTCGTGTTGAAGGAGATGAAGATGAGGATGATGTTGATGATATCGAGAATGAGTTCA--- 484
Qy       487 GGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACG 546
Db       485 ---------ATTACGCCCAGGGAGCTAACAAGGCGAGACACCAACGCCA---TGGCGAAG 532
Qy       547 GCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCAC 606
Db       533 AGTTTTCTTCTTCCTCTAGACATGAATCTCAACCAATTCCTCTTCTCACCCATGGCCATA 592
Qy       607 TCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTT 666
Db       593 CGGTTTCTGGAGAGATTCGCACGCCTGATACACAATCTGTGCGAACTACATCAGGTCCTT 652
Qy       667 CTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTAC 726
Db       653 ------TGGGTCCTTCTGACAGGAATGCTATTTCATCTCCATATATTGATCCACGGCAAC 706
Qy       727 CTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTG 786
Db       707 CTGTCCCTGTAAGAATCGTGGACCCGTCAAAAGACTTGAACTCTTATGGGCTTGGTAATG 766
Qy       787 TTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATG---------- 836
Db       767 TTGACTGGAAAGAAAGAGTTGAAGGCTGGAAGCTGAAGCAGGAGAAAAATATGTTACAGA 826
Qy       837 ----------------CACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATG 879
Db       827 TGACTGGTAAATACCATGAAGGGAAAGGAGGAGAAATTGAAGGGACTGGTTCCAATGGCG 886
Qy       880 CTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCAT 939
Db       887 AAGAACTCCAAATGGCTGATGATACACGTCTTCCTATGAGTCGTGTGGTGCCTATCCCAT 946
Qy       940 CAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCT 999
Db       947 CTTCTCGCCTAACCCCTTATCGGGTTGTGATTATTCTCCGGCTTATCATCTTGTGTTTCT 1006
Qy      1000 TCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTG 1059
Db      1007 TCTTGCAATATCGTACAACTCACCCTGTGAAAAATGCATATCCTTTGTGGTTGACCTCGG 1066
Qy      1060 TTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCC 1119
Db      1067 TTATCTGTGAGATCTGGTTTGCATTTTCTTGGCTTCTTGATCAGTTTCCCAAATGGTACC 1126
Qy      1120 CTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGC 1179
Db      1127 CCATTAACAGGGAGACTTATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAAC 1186
Qy      1180 CATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTC 1239
Db      1187 CATCACAGCTCGTTCCTGTTGATGTGTTTGTTAGTACAGTGGACCCATTGAAAGAGCCTC 1246
Qy      1240 CTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGG 1299
Db      1247 CCCTTGTTACAGCAAACACAGTTCTCTCGATTCTTTCTGTGGACTACCCGGTAGATAAAG 1306
Qy      1300 TTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAA 1359
```

```
Db      1307 TAGCCTGTTATGTTTCAGATGATGGTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAA 1366
Qy      1360 CATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCG 1419
Db      1367 CCGCTGAGTTTGCAAAGAAATGGGTACCATTTTGCAAGAAATTCAACATTGAACCTAGGG 1426
Qy      1420 CTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACT 1479
Db      1427 CCCCTGAATTCTATTTTGCCCAGAAGATAGATTACTTGAAGGACAAGATCCAACCGTCTT 1486
Qy      1480 TTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATG 1539
Db      1487 TTGTTAAAGAGCGACGAGCTATGAAGAGAGAGTATGAAGAGTTTAAAGTGAGGATAAATG 1546
Qy      1540 CCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCC 1599
Db      1547 CTCTTGTTGCCAAAGCACAGAAAATCCCTGAAGAAGGCTGGACAATGCAGGATGGTACTC 1606
Qy      1600 CCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAA 1659
Db      1607 CCTGGCCTGGTAACAACACTAGAGATCATCCTGGAATGATACAGGTGTTCTTAGGCCATA 1666
Qy      1660 GCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGA 1719
Db      1667 GTGGGGGTCTGGATACCGATGGAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAA 1726
Qy      1720 AACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCT 1779
Db      1727 AGCGGCCTGGATTTCAACACCACAAAAAGGCTGGAGCTATGAATGCATTGATCCGTGTAT 1786
Qy      1780 CTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACA 1839
Db      1787 CTGCTGTTCTTACCAATGGAGCATATCTTTTGAACGTGGATTGTGATCATTACTTTAATA 1846
Qy      1840 ACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGG 1899
Db      1847 ACAGTAAGGCTATTAAAGAAGCTATGTGTTTCATGATGGACCCGGCTATTGGAAAGAAGT 1906
Qy      1900 TTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTA 1959
Db      1907 GCTGCTATGTCCAGTTCCCTCAACGTTTTGACGGTATTGATTTGCACGATCGATATGCCA 1966
Qy      1960 ACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAA 2019
Db      1967 ACAGGAATATAGTCTTTTTCGATATTAACATGAAGGGGTTGGATGGTATCCAGGGTCCAG 2026
Qy      2020 TTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCA 2079
Db      2027 TATATGTGGGTACTGGTTGTTGTTTTAATAGGCAGGCTCTATATGGGTATGATCCTGTTT 2086
Qy      2080 AAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTT 2139
Db      2087 TGACGGAAGAAGATTTAGAACCAAATATTA----------TTGTCAAGA 2125
Qy      2140 GCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGT 2199
Db      2126 GCTGTTGCGGGTCAAGGAAGAAAGGTAAAAGTAGCAA---GAAGTATAACTACGAAAAGA 2182
Qy      2200 TATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACG 2259
Db      2183 GGAGAGGCATCAACAGAAGTGACTCCAATGCTCCACTTTTCAATATGGAGGACATCGATG 2242
Qy      2260 AAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAA 2319
Db      2243 AGGGTTTTGAAGGTTATGATGATGAGAGGTCTATTCTAATGTCCCAGAGGAGTGTAGAGA 2302
Qy      2320 AGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCT 2379
Db      2303 AGCGTTTTGGTCAGTCGCCGGTATTTATTGCGGCAACCTTCATGGAACAAGGCGGCATTC 2362
Qy      2380 TGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTT 2439
Db      2363 CACCAACAACCAATCCCGCTACTCTTCTGAAGGAGGCTATTCATGTTATAAGCTGTGGTT 2422
```

```
Qy       2440 ATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAG 2499
             | ||||||||||| || ||||| |||||||||| |||||||||| || || || ||||
Db       2423 ACGAAGACAAGACTGAATGGGGCAAAGAGATTGGTTGGATCTATGGTTCCGTGACGGAAG 2482

Qy       2500 ATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTA 2559
             ||||||| ||||| |||||||||||    | ||||||   || || |||||||  |||
Db       2483 ATATTCTTACTGGGTTCAAGATGCATGCCCGGGGTTGGATATCGATCTACTGCAATCCTC 2542

Qy       2560 AACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGC 2619
               |||   ||| ||||| || |||||||   | |||||||| |||||| |  |||| || |
Db       2543 CACGCCCTGCGTTCAAGGGATCTGCACCAATCAATCTTTCTGATCGTTTGAACCAAGTTC 2602

Qy       2620 TTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATG 2679
             |||| |||||| | || ||||| |||||  |  | ||||  ||||| ||| | |||||||
Db       2603 TTCGATGGGCTTTGGGATCTATCGAGATTCTTCTTAGCAGACATTGTCCTATCTGGTATG 2662

Qy       2680 GGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATC 2739
             | ||   |||  |  ||| | ||||||| ||   |  | || |||||| ||||||| ||||
Db       2663 GTTACCATGGAAGGTTGAGACTTTTGGAGAGGATCGCTTATATCAACACCATCGTCTATC 2722

Qy       2740 CTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAG 2799
             ||   ||||| || || ||  | || || ||||   | || ||  | ||| |  | || |
Db       2723 CTATTACATCCATCCCTCTTATTGCGTATTGTATTCTTCCCGCTTTTTGTCTCATCACCG 2782

Qy       2800 GGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTA 2859
               | ||| ||||  || ||| | |  ||    || ||  | ||||||||   ||| || |
Db       2783 ACAGATTCATCATACCCGAGATAAGCAACTACGCGAGTATTTGGTTCATTCTACTCTTCA 2842

Qy       2860 TCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACT 2919
             |||  |||  ||  ||  | ||||| ||  |||||||||| |||||  | |||||  || |
Db       2843 TCTCAATTGCTGTGACTGGAATCCTGGAGCTGAGATGGAGCGGTGTGAGCATTGAGGATT 2902

Qy       2920 GGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGT 2979
             |||||||||| |||||||||||||||||||| ||    ||| | || || ||||||||  |
Db       2903 GGTGGAGGAACGAGCAGTTCTGGGTCATTGGTGGCACATCCGCCCATCTTTTTGCTGTCT 2962

Qy       2980 TCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGG 3039
             |||| || || || |||||  | |||||| | || || | |||||||||| ||||| || |
Db       2963 TCCAAGGTCTACTTAAGGTTCTTGCTGGTATCGACACCAACTTCACCGTTACATCTAAAG 3022

Qy       3040 ---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
                |  || |||| ||| ||  |||| || || |  ||||||||||  |  |  | ||
Db       3023 CCACAGACGAAGATGGGGATTTTGCAGAACTCTACATCTTCAAATGGACAGCTCTTCTCA 3082

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
             | || || |||||| | || ||  ||||| |||| ||  | || ||||| || ||  |||
Db       3083 TTCCACCAACCACCGTCCTACTTGTGAACCTCATAGGCATTGTGGCTGGTGTCTCTTATG 3142

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
             |  | || |  || ||  |||| ||||| || || || |||||||| ||||| || || |
Db       3143 CTGTAAACAGTGGCTACCAGTCGTGGGGTCCGCTTTTCGGGAAGCTCTTCTTCGCCTTAT 3202

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
             |||| |||| |||||| || || ||  | |||||| || | ||||| |||||| | ||||
Db       3203 GGGTTATTGCCCATCTCTACCCTTTCTTGAAAGGTCTGTTGGGAAGACAAAACCGAACAC 3262

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
             |||| || ||||| |||||||| |||| || || || |||||||| | ||||||||
Db       3263 CAACCATCGTCATTGTCTGGTCTGTTCTTCTCGCCTCCATCTTCTCGTTGCTTTGGGTCA 3322

Qy       3337 GGATTGATCCTTTCCTTG 3354
             ||||  |||| ||  | |
Db       3323 GGATCAATCCCTTTGTGG 3340
```

Search completed: December 15, 2003, 22:42:12
Job time : 14024 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 17:49:46 ; Search time 7630 Seconds
(without alignments)
12537.669 Million cell updates/sec

Title: US-09-720-383C-9
Perfect score: 3936
Sequence: 1 cttctccctcgtcggtgcgg.....ctgagttggttttgtgaaaa 3936

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3869.8	98.3	3897	11	AY103655	AY103655 Zea mays
2	2591	65.8	3783	11	AY104730	AY104730 Zea mays
3	2225.2	56.5	3788	11	AY103701	AY103701 Zea mays
4	1342.8	34.1	3898	11	AY110415	AY110415 Zea mays
5	1289.4	32.8	3763	11	AY108113	AY108113 Zea mays
6	1212.4	30.8	3696	11	AY110079	AY110079 Zea mays
7	1209.6	30.7	3728	11	AY112236	AY112236 Zea mays
8	1161.6	29.5	2872	11	AY104236	AY104236 Zea mays
c 9	811	20.6	877	29	CC379931	CC379931 PUHOP50TB
c 10	731.4	18.6	758	14	CD441181	CD441181 EL01N0553
11	704.6	17.9	856	14	CB633016	CB633016 OSIIEb11J
12	686.2	17.4	859	14	CB645001	CB645001 OSJNEb06H
13	683.2	17.4	864	14	CB627056	CB627056 OSIIEb01L
14	672.4	17.1	862	14	CD437450	CD437450 EL01N0501
c 15	670.6	17.0	900	29	CC365622	CC365622 PUHBR52TD
16	652	16.6	792	14	CB671068	CB671068 OSJNEe04H
17	641.2	16.3	783	14	CB671503	CB671503 OSJNEe05B
18	626.6	15.9	1022	29	BZ827769	BZ827769 PUFHO17TB
19	621.8	15.8	795	14	CB677472	CB677472 OSJNEe14G
20	612.8	15.6	806	14	CB629322	CB629322 OSIIEb05H
21	596.4	15.2	758	14	CB672249	CB672249 OSJNEe06D
22	594.6	15.1	714	14	CB629928	CB629928 OSIIEb06I
23	580.6	14.8	657	9	AW037064	AW037064 614025D05
24	578.6	14.7	765	14	CB676377	CB676377 OSJNEe12J
25	570.6	14.5	670	29	CC160666	CC160666 ig10c05.g
26	566.6	14.4	827	14	CB677498	CB677498 OSJNEe14I
27	549	13.9	701	13	BQ803259	BQ803259 WHE2835_E
28	546.8	13.9	770	14	CB681901	CB681901 OSJNEf08K
29	539.6	13.7	789	14	CB672865	CB672865 OSJNEe07C
30	539	13.7	591	14	CD425359	CD425359 SA1_11_F0
c 31	537.8	13.7	930	29	CC354726	CC354726 PUHLW20TB
32	537.6	13.7	635	9	AW146609	AW146609 614074E01
c 33	533.4	13.6	642	9	AW053172	AW053172 614072C08
c 34	532	13.5	660	9	AI621823	AI621823 486027A04
35	525.4	13.3	539	13	BU197567	BU197567 946162C06
36	524.8	13.3	745	14	CD444028	CD444028 EL01N0434
37	524.4	13.3	695	13	BU100946	BU100946 WHE3359_G
38	522.6	13.3	755	13	BQ484013	BQ484013 WHE3515_D
39	517.2	13.1	760	10	BE412810	BE412810 MCG008.D0
40	517	13.1	693	10	BF628750	BF628750 HVSMEb000
41	515	13.1	700	9	AL503528	AL503528 AL503528
42	510	13.0	510	9	AW126453	AW126453 614072C08
c 43	505.4	12.8	515	9	AI920493	AI920493 614025D05
44	502.8	12.8	731	13	BQ294855	BQ294855 WHE2855_B
45	501.8	12.7	619	14	CA019440	CA019440 HV11P23r

ALIGNMENTS

RESULT 1
AY103655
LOCUS AY103655 3897 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO096398 mRNA sequence.
ACCESSION AY103655
VERSION AY103655.1 GI:21206733
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3897)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3897)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES Location/Qualifiers
source 1. .3897
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635496"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 982 a 839 c 1008 g 1068 t

ORIGIN

Query Match 98.3%; Score 3869.8; DB 11; Length 3897;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3882; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         14 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTC-GGGG 72

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 132

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 192

Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 252

Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 312

Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 372

Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 432

Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 492

Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 552

Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 612

Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 672

Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 732

Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 792

Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 852

Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 912

Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        913 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 972

Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1032

Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1033 ATCCGGTGAATGATGCATTTGCTTTGTGGCGCATATCTGTTATCTGTGAAATCTGGTTTG 1092

Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1152

Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1212

Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1272

Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1332

Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1392

Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1393 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1452

Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1453 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1512

Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1513 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1572

Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1573 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1632

Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1633 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1692

Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1693 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1752

Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1753 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1812

Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1813 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1872
Qy      1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db      1873 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1932
Qy      1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db      1933 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1992
Qy      1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db      1993 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2052
Qy      2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db      2053 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2112
Qy      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db      2113 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2172
Qy      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db      2173 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2232
Qy      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db      2233 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2292
Qy      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Db      2293 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2352
Qy      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Db      2353 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2412
Qy      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Db      2413 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2472
Qy      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Db      2473 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2532
Qy      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Db      2533 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2592
Qy      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Db      2593 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2652
Qy      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Db      2653 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2712
Qy      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Db      2713 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2772
Qy      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Db      2773 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2832
Qy      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Db      2833 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2892
Qy      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Db      2893 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2952
```

```
Qy      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Db      2953 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3012
Qy      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Db      3013 TAGCTGGTGATGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3072
Qy      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Db      3073 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3132
Qy      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Db      3133 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3192
Qy      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Db      3193 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3252
Qy      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Db      3253 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3312
Qy      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
Db      3313 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3372
Qy      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
Db      3373 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3432
Qy      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
Db      3433 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3492
Qy      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
Db      3493 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3552
Qy      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
Db      3553 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3612
Qy      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
Db      3613 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3672
Qy      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
Db      3673 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3732
Qy      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
Db      3733 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3792
Qy      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
Db      3793 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3852
Qy      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAAT 3885
Db      3853 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAAT 3897
```

```
RESULT 2
AY104730
LOCUS       AY104730               3783 bp    mRNA    linear   HTC 16-OCT-2002
DEFINITION  Zea mays PCO100501 mRNA sequence.
ACCESSION   AY104730
VERSION     AY104730.1  GI:21207808
KEYWORDS    HTC.
```

```
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3783)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 3783)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1. .3783
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:635765"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      933 a    908 c   1002 g    940 t
ORIGIN

  Query Match             65.8%;  Score 2591;  DB 11;  Length 3783;
  Best Local Similarity   86.5%;  Pred. No. 0;
  Matches 2884;  Conservative    0;  Mismatches  440;  Indels    9;  Gaps    2;

Qy        104 CTCTGCCTTCGTCGGCGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCT 163
              | ||   |  ||     | |  ||  |||  | |      | ||||||||| ||||||||||
Db        146 CGCTTGCCCCGGGATCTGGAGATCTGGTAGCGCCAGGGGGATGGAGGCCAGCGCCGGGCT 205

Qy        164 GGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGG 223
              ||| |||||||| |||||||| |||||||||||||||||||||||||||| ||||| || ||
Db        206 GGTCGCCGGCTCGCACAACCGGAACGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGG 265

Qy        224 GCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGG 283
              |||||||||    |    ||| |||||| ||||||||||||||||||||||||||||| ||
Db        266 GCCGAAGCCCATGGACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGG 325

Qy        284 CCTTGCCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCG 343
              |    |||||  |||||| || |||||||| ||||||||||||||||||||||||| |||||||
Db        326 GCGCAACCCCGACGGGGAGCCGTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCG 385

Qy        344 GGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCG 403
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||||||| ||
Db        386 GGACTGCTACGAGTACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCG 445

Qy        404 ATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGA 463
              | |||||||||||||| |||    || ||| |||| ||||||||||||||||||||||||
Db        446 CTTCAAGCGCCTCAAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGA 505

Qy        464 TGACCTGGACAACGAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTC 520
              |||||||| ||||||||||||||    |||    || ||||| ||||    | ||||||||
Db        506 CGACCTGGAGAACGAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTC 565

Qy        521 CATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGC 580

Db        566 CATGCTCCACGCCCACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCC 625

Qy        581 TTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCC 640
              |||||  | || |||||||||||| |||||||||||||| || ||||| ||||||||||||
Db        626 ATTCCACCCCATCCCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCC 685

Qy        641 ACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGGAAAGAGGATACATCC 700
              ||||| |||||||| || ||||| || ||| ||||||| ||||| |||||||| || ||
Db        686 GCCGGACCAGCACGCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGGAAGAGGATTCACCC 745

Qy        701 CCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGA 760
              || || || |||||||||| | | |||||||||| |||||||||||||| ||||||||
Db        746 TCTCCCGTACGCGGATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGA 805

Qy        761 TCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCA 820
              ||| || |||||||| || || || || || |||||||| ||||||||| |||||||||
Db        806 TCTCGCCGCATATGGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCA 865

Qy        821 GAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGC 880
              ||  || |||||||||||||||||| |||| ||||| || || |||||||||| |||||
Db        866 GAAGCAGGAGAGGATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGC 925

Qy        881 TGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATC 940
              |||||||||||||||||||||||| ||||| | | ||||||| || || || ||||| ||
Db        926 AGATCTACCACTAATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTC 985

Qy        941 AAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTT 1000
              |||||| || || || |||||||||||||| || |||||||| ||||||||| | |||||
Db        986 AAGCCAAATCAACCCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTT 1045

Qy       1001 CTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGT 1060
              |||||||||||||||||||||||||||||  |||||||||||||| ||||||||||||||
Db       1046 CTTCCACTACCGAGTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGT 1105

Qy       1061 TATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCC 1120
              |||||||||| |||||||||||||||||||||||||| || || |||||||||| ||
Db       1106 GATCTGTGAAATTTGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCC 1165

Qy       1121 TATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCC 1180
              ||| ||||| || || || |||||||||     | ||||| |||||||||| || ||
Db       1166 TATCGAGAGGGAAACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCC 1225

Qy       1181 ATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCC 1240
              |||||||| || ||  | |||||||||||||||||||||||||||| ||||||||||
Db       1226 TTCTCAACTCGCCCCTGTCGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCC 1285

Qy       1241 TTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGT 1300
              |||||||| | |||||||||||||||||||||||||||||||||| ||||||||||
Db       1286 ATTGGTCACTGCTAATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGT 1345

Qy       1301 TTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAAC 1360
              ||| ||||| |||||||||||||||||||| ||||| || |||||||||| |||||||||
Db       1346 TTCATGCTACGTTTCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAAC 1405

Qy       1361 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGC 1420
              |||||||||||||||||||||||||||||||||||||| | || |  |||| ||||| ||
Db       1406 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGC 1465

Qy       1421 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTT 1480
              |||||||||||||||||||||||||||||||||| ||||||||||||||  |||||||
Db       1466 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTT 1525

Qy       1481 TGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGC 1540
              |||||| ||  |||||||||||||||||||||||||||||||||||| ||||||||||
Db       1526 TGTTAGAGAACGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGC 1585

Qy       1541 CTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCC 1600
              ||| ||||| |||||||| || |||||||| ||||||||||||||||| |||||||| ||
Db       1586 CTTGGTTGCTAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCC 1645

Qy       1601 CTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAG 1660
              ||||| ||||| ||||| |||||||||||||||||||||||| |||||||| |||||
```

```
Db       1646 ATGGCCCGGAAATAATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAG 1705

Qy       1661 CGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAA 1720

Db       1706 TGGTGGCCATGATGTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAA 1765

Qy       1721 ACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1780

Db       1766 ACGGCCAGGCTACAACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1825

Qy       1781 TGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAA 1840

Db       1826 TGCTGTACTAACTAATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAA 1885

Qy       1841 CAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGT 1900

Db       1886 TAGTAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAATAAAGT 1945

Qy       1901 TTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAA 1960

Db       1946 TTGCTATGTGCAGTTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAA 2005

Qy       1961 CCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAAT 2020

Db       2006 CAGAAATGTTGTCTTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAAT 2065

Qy       2021 TTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAA 2080

Db       2066 TTATGTGGGTACTGGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAA 2125

Qy       2081 AACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTG 2140

Db       2126 AACAAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTG 2185

Qy       2141 CTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAA 2194

Db       2186 CTGTTTTGGTAACAGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAA 2245

Qy       2195 AAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAAT 2254

Db       2246 GATAAAGAAACTTTTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAAT 2305

Qy       2255 TGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATT 2314

Db       2306 TGATGAAGCTGCTCCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTT 2365

Qy       2315 AGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGG 2374

Db       2366 GGAAAAGAAATTTGGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGG 2425

Qy       2375 AACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTG 2434

Db       2426 AACCCTGAAGAGTGCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGCTG 2485

Qy       2435 TGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTAC 2494

Db       2486 TGGATATGAAGACAAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCAC 2545

Qy       2495 AGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCAT 2554

Db       2546 AGAAGATATTCTTACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCAT 2605

Qy       2555 ACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCA 2614

Db       2606 ACCTAAACGGGCCGCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTCTTCACCA 2665

Qy       2615 GGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTG 2674

Db       2666 GGTTCTTCGGTGGGCTCTTGGTTCAATTGAAATTTTCTTCAGCAACCACTGCCCTCTCTG 2725

Qy       2675 GTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGT 2734

Db       2726 GTATGGGTATGGTGGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGT 2785

Qy       2735 GTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATT 2794

Db       2786 ATACCCTTGGACATCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCT 2845

Qy       2795 GACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACT 2854

Db       2846 GACAGGGAAATTTATCACGCCAGAGCTTAACAATGTTGCCAGCCTGTGGTTCATGTCACT 2905

Qy       2855 TTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGA 2914

Db       2906 TTTCATCTGCATTTTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGA 2965

Qy       2915 TGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGC 2974

Db       2966 TGACTGGTGGAGAAACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGC 3025

Qy       2975 TGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATC 3034

Db       3026 TGTGTTCCAGGGACTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATC 3085

Qy       3035 AAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATT 3094

Db       3086 CAAGGGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCT 3145

Qy       3095 GATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAA 3154

Db       3146 GATACCTCCGACAACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAA 3205

Qy       3155 TGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATT 3214

Db       3206 TGCGATCAACAACGGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATT 3265

Qy       3215 TTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGAC 3274

Db       3266 TTGGGTGATCGTCCATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGAC 3325

Qy       3275 ACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGT 3334

Db       3326 GCCAACGATTGTCATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGT 3385

Qy       3335 TCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGA 3394

Db       3386 CCGGATCGACCCGTTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGA 3445

Qy       3395 TTGCAACTAGGATGTCAGTGCATCAGCTCCCCC 3427

Db       3446 TTGCAACTAGGAGGTCAGCACGTGGACTTCCCC 3478
```

RESULT 3
AY103701

LOCUS	AY103701	3788 bp mRNA linear HTC 16-OCT-2002
DEFINITION	Zea mays PCO120363 mRNA sequence.	
ACCESSION	AY103701	
VERSION	AY103701.1 GI:21206779	
KEYWORDS	HTC.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
REFERENCE	1 (bases 1 to 3788)	
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.	
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes	
JOURNAL	Unpublished (2002)	
REFERENCE	2 (bases 1 to 3788)	
AUTHORS	Coe,E.H.	
TITLE	Direct Submission	
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA	

```
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1. .3788
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:637005"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      851 a    954 c   1114 g    866 t      3 others
ORIGIN

  Query Match             56.5%;  Score 2225.2;  DB 11;  Length 3788;
  Best Local Similarity   81.1%;  Pred. No. 0;
  Matches 2685;  Conservative    0;  Mismatches  586;  Indels   40;  Gaps    7;

Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| |||||||||||||||||| || |||
Db        180 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 239

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || |||||||| || || |||||||||||| |||| ||  | ||||  | || ||  |
Db        240 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 299

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||     |||||       | |||||||||| |||||  | | |  || |
Db        300 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 359

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||||||| | |||||||||
Db        360 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 419

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||||| |||||| |||||    |||||  |||||||  ||| ||  ||||||||||||
Db        420 TACGAGCGCCGCGAGGGCTCGCAGGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 479

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              |||||  ||||   ||  |||  |  || |||||||||||||||||||||||||  ||||||||    |
Db        480 AAGGGTTGCCCGCGGGTGGCTGGGGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 539

Qy        477 GAGTT---------------CAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              |||||                |  |  |||   | |||  ||||||   || ||||||||||
Db        540 GAGTTCGGCCTGCAGGACGGCGCCGGCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 599

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
              |||||    |  || |||||||||||||| || || ||          |||||  |  |
Db        600 ATGCTGCGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCGGAC 650

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
              |||  |  |  |||||| || || |||||||||||||| || |||||  ||||||||||
Db        651 TTCAACCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 710

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG------GGTGGTGGGGGAAAGAGGATA 695
              |||||||||||||||| ||||| || |  ||||      || || || || ||||||||
Db        711 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGGGCGGCGGCGGCGGCGGCAAGAGGATC 770

Qy        696 CATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCC 755
              || || || || |  || ||| ||| | | || |||||||| || || |||||||| |||
Db        771 CACCCGCTCCCCTTCGCTGATTCCAACCTTCCAGTGCAACCGAGATCCATGGACCCGTCC 830

Qy        756 AAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGG 815
              |||||||| |||||||| ||||| || || |||||||| |||||||| |||||  | ||
Db        831 AAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGCTGG 890

Qy        816 AAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GATGGT 872
              |||||||  || ||| |  |||| ||     |||  || ||||| ||||||   |||||
Db        891 AAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGATGGC 950

Qy        873 GACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCA 932
              |||||||| ||||| |||||||||||||||||| || || | | ||||||| ||| | ||
Db        951 GACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTCCCT 1010

Qy        933 CTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTG 992
              |  |||||||||  |||||||| || |||||||||||| |||| ||| | |||||||||
Db       1011 ATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTTTTG 1070

Qy        993 GGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTC 1052
              || |||||||||||||||||||||||||||||||||||||| ||| ||||||||||| ||||||||||
Db       1071 GGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGGCTC 1130

Qy       1053 ATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAG 1112
              |||||||| |||||||||||||||||||| |||||  |||||||||||||| |||||||||
Db       1131 ATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCAAAG 1190

Qy       1113 TGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAA 1172
              |||  |  ||  || |||||||||||||||||| |||||  |||||||  |||||  |||||||||
Db       1191 TGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAGGAA 1250

Qy       1173 GGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAG 1232
              || ||  ||  |||||  |||||||||||  || ||||||||||||||||||||||  ||||
Db       1251 GGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACAAAG 1310

Qy       1233 GAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTT 1292
              |||||||| |||||||||  | ||  |||||  ||  || ||||||||  |||||||||||  |||
Db       1311 GAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCGGTT 1370

Qy       1293 GATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTA 1352
              ||  |||||  || ||||||||||||||||||||||||||||||||  ||||||||||||||
Db       1371 GAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCATTG 1430

Qy       1353 TCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAA 1412
              |||||||||||||||||||||||||||||||||||||||| |||||  ||  |||||  ||
Db       1431 TCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATCGAG 1490

Qy       1413 CCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCA 1472
              |||||  |||||  ||||||||||||||||||||||||||||||  |||||||||||||  ||
Db       1491 CCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTTGCT 1550

Qy       1473 GCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGA 1532
              ||    |||||||||||||||  || ||||||||||  || ||||||||||||||  ||
Db       1551 GCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTAAGG 1610

Qy       1533 ATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGAT 1592
              ||||||||||  |||||  ||||||||  || ||||||||  ||||||||||||||||||||
Db       1611 ATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAAGAT 1670

Qy       1593 GGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTT 1652
              ||||  ||||||||||||||||||  ||  ||||||||||||||||||||||  ||||||
Db       1671 GGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTCCTT 1730

Qy       1653 GGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCT 1712
              ||||||||  ||  ||  |  |||     ||  ||||||||  ||||  ||  ||||||||||  ||
Db       1731 GGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTCTCG 1790

Qy       1713 AGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTC 1772
              |||||  ||  |  |||||  ||||||||||  |||||  ||||||||  |||||||||  |||||
Db       1791 AGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTGGTC 1850

Qy       1773 CGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTAC 1832
              ||  ||||||||||   ||  ||||||||  ||||  |  ||  ||||||||  ||||||||||||
Db       1851 CGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCACTAC 1910

Qy       1833 ATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGA 1892
              ||||||||  ||||||||  |||||  ||  ||  ||||||||  ||||||||  |||||  |  ||
```

```
Db      1911 ATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTGGGG 1970

Qy      1893 AAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGA 1952
             ||||| || |||||||||||||||||||| || |||||||| |||||      |||| |||
Db      1971 AAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGATCGA 2030

Qy      1953 TATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAG 2012
             || |||||| |||| |||||||||||||| |||||||||||||||||||| ||||||||
Db      2031 TACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATTCAA 2090

Qy      2013 GGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGAT 2072
             || || |||||||| |||||||||||||| || ||| |||||||| | ||||||||||||
Db      2091 GGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTATGAT 2150

Qy      2073 GCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTT 2132
             || || ||||| |||||||||||||||||| |||||||||||||||||||||||||||| |
Db      2151 GCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGCCTC 2210

Qy      2133 TGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAG 2192
             |  || ||||||    | || | |||  |||||||||||| ||||| || || ||||||
Db      2211 TCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAGAAG 2270

Qy      2193 AAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAA 2252
             || ||   |||||||||||||||||| ||| |||| ||| |||||||||||| | |||||||
Db      2271 AAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGTGAA 2330

Qy      2253 ATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAA 2312
             ||||| |||| ||||||||| ||||| |  || |||||||| || |||||||||| |||
Db      2331 ATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAGAAA 2390

Qy      2313 TTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGT 2372
              |||| |||||||||| || |||||||||||||   |||| |||||||| ||||| ||
Db      2391 CTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAACGGA 2450

Qy      2373 GGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGT 2432
             || ||| |||||||| |||||||| |||||||||| |||| |||||||||||||| || ||
Db      2451 GGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATCAGC 2510

Qy      2433 TGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTT 2492
             || || || |||||||||| |||||||||||||||||||||||||| || |||||  |
Db      2511 TGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCGATC 2570

Qy      2493 ACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGC 2552
             ||||| |||||  | ||||| || |||||||| || |||||| |||||||| |||||||||
Db      2571 ACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTACTGC 2630

Qy      2553 ATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAC 2612
             || || || |||  |||||||||||||||||| |||||||| |||||| || |||||||||
Db      2631 ATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTTCAC 2690

Qy      2613 CAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC-TTCAGCAATCATTGCCCTCT 2671
             ||||| ||||| ||||| |||||||||  | || || ||| |||||  || || ||   | ||
Db      2691 CAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTTCAGNAANCAATGGNCACT 2750

Qy      2672 TTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCAT 2731
             |||||| || || || ||||| || |||||  |||||| ||||| || ||||||||||
Db      2751 TTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCCAT 2810

Qy      2732 CGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTT 2791
             |||  ||||| ||||| || ||||| ||| |||||||||||| ||||||||||||||   |
Db      2811 CGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGCCT 2870

Qy      2792 ATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTC 2851
               | || ||||| |||||||| |||||||| |  ||||| |||||  | ||||||||| |
Db      2871 GCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATGGC 2930

Qy      2852 ACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAAT 2911
             |||||| |||||||| |   |  ||  |||||||| |||||||| |||||||| || |  ||
Db      2931 ACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCCAT 2990

Qy      2912 TGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTT 2971
              || |||||||||||||| |||||||||||||||| ||||| || ||  | || || ||
Db      2991 CGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTGTT 3050

Qy      2972 TGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGAC 3031
             || |||||||||||| || || |||||  | || ||  | || || ||||||||||||||
Db      3051 CGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTGAC 3110

Qy      3032 ATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTT 3091
             || |||||  || || || |||||||||||| |||||  || || |||||  ||||  |||
Db      3111 GTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACCCT 3170

Qy      3092 ATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTC 3151
             |||||||  || |||||  | || ||  |||||||||| || |||||  || ||  | ||
Db      3171 GCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATCTC 3230

Qy      3152 AAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGC 3211
             || |||||||| |||||  || |||||  |||||||||||  || |||||||| |||||  ||
Db      3231 GAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTCGC 3290

Qy      3212 ATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAG 3271
             || |||||||| |||||  || || || || |||||  |||  ||||  ||  |||||  |||||
Db      3291 GTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAACAG 3350

Qy      3272 GACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTG 3331
             |||  || ||||| ||||||||||||||||  |||||||| ||  ||||||||||||||  ||
Db      3351 GACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTGTG 3410

Qy      3332 GGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTT 3391
             |||  ||  |  || || ||||| || |||   || || |||||  || ||||||||||  |
Db      3411 GGTCCGCGTCGACCCGTTCCTCGCCAAGAGCGACGGCCCGCTCCTGGAGGAGTGTGGCCT 3470

Qy      3392 GGATTGCAACT 3402
             |||  |||||||
Db      3471 GGACTGCAACT 3481
```

RESULT 4
AY110415

```
LOCUS       AY110415                3898 bp    mRNA    linear   HTC 17-OCT-2002
DEFINITION  Zea mays CL1166_1 mRNA sequence.
ACCESSION   AY110415
VERSION     AY110415.1  GI:21214824
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3898)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 3898)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1. .3898
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:630057"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
```

assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 932 a 841 c 991 g 1054 t 80 others
ORIGIN

Query Match 34.1%; Score 1342.8; DB 11; Length 3898;
Best Local Similarity 64.9%; Pred. No. 1.3e-289;
Matches 2120; Conservative 0; Mismatches 1034; Indels 111; Gaps 8;

```
Qy   127 GTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCA 186
Db   172 GTAGGAGGAGCGCGGAGATGGCGGCCAACAAGGGGATGGTGGCGGGCTCGCACAACCGCA 231

Qy   187 ACGAGCTCGTCGTCATCCGCCGCGACGGCGATCC------CGGGCCGAAGCCGCCGCGGG 240
Db   232 ACGAGTTCGTCATGATCCGCCACGACGGCGACGCGCCGGGCCCGGCTAAGCCCACGAAGA 291

Qy   241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
Db   292 GTGCGAATGGGCAGGTCTGCCAGATTTGCGGCGACTCTGTGGGCGTTTCAGCCACTGGTG 351

Qy   301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
Db   352 ATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCTTGCTATGAGTACG 411

Qy   361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
Db   412 AGCGCAAGGAGGGGAACCAATGCTGCCCTCAGTGCAAGACTAGATACAAGAGACAGAAAG 471

Qy   421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Db   472 GTAGCCCTCGAGTTCATGGTGATGAGGATGAGGAAGATGTTGATGACCTAGACAATGAAT 531

Qy   481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Db   532 TCAACTACAA----------------------GCAAGGCAGTGGGAAGGGCCCAGAGT 567

Qy   541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAAT- 599
Db   568 GGCAGCTGCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCCACACCATC 627

Qy   600 --GTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGC 657
Db   628 GGATTCCACGCCTGACAAGCGGTCAACAGATATCTGGAGAGATCCCTGATGCATCCCCTG 687

Qy   658 TGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATC 717
Db   688 ACCGTCATTCTATC----------------CGCAGTCCAACATCGAGCTATGTTGATC 729

Qy   718 CCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGT 777
Db   730 CAAGCGTCCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTCCTATGGGC 789

Qy   778 ATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGC 837
Db   790 TTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGACNNNNNTA 849

Qy   838 ACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG---------------------- 873
Db   850 TGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGGAGACATGGAGGGGACTGGCT 909

Qy   874 ---ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTC 930
Db   910 CAAATGGAGAAGATATGCAAATGGTTGATGATGCACGCCTACCTTTGAGCCGCATCGTGC 969

Qy   931 CACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTT 990
Db   970 CAATTTCCTCAAACCAGCTCAACCTTTACCGGGTAGTGATCATTCTCCGTCTTATCATCC 1029

Qy   991 TGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGC 1050

Db  1030 TGTGCTTCTTCTTCCAGTATCGTATCAGTCATCCAGTGCGTGATGCTTATGGATTGTGGC 1089

Qy  1051 TCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAA 1110
Db  1090 TAGTATCTGTTATCTGCGAGGTCTGGTTTGCCTTGTCCTGGCTTCTAGATCAGTTCCCAA 1149

Qy  1111 AGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGG 1170
Db  1150 AATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTGAGGTATGATAGAG 1209

Qy  1171 AAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAA 1230
Db  1210 AGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTCGTCAGTACAGTGGATCCATTGA 1269

Qy  1231 AGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTG 1290
Db  1270 AGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTGCTGTGGATTACCCTG 1329

Qy  1291 TTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCAT 1350
Db  1330 TTGACAAAGTGTCATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGAGTCTC 1389

Qy  1351 TATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTG 1410
Db  1390 TCTCTGAAACCGCCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGAAGCACAATATTG 1449

Qy  1411 AACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGG 1470
Db  1450 AACCAAGAGCTCCAGAATTTTACTTTGCTCNNNNNNTAGATTACCTGAAGGACAAAATTC 1509

Qy  1471 CAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGA 1530
Db  1510 AACCTTCATTTGTTAAGGAAAGACGCGCAATGAAGAGGGAGTATGAAGAATTCAAAGTAA 1569

Qy  1531 GAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAG 1590
Db  1570 GAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCATGGCTG 1629

Qy  1591 ATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCC 1650
Db  1630 ATGGAACTGCTTGGCCTGGGAATAACCCTAGGGACCATCCTGGCATGATTCAGGTGTTCT 1689

Qy  1651 TTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTT 1710
Db  1690 TGGGGCACAGTGGTGGGCTCGACACTGATGGAAATGAGTTACCACGTCTTGTCTATGTCT 1749

Qy  1711 CTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGG 1770
Db  1750 CTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAATGAATGCGCTGA 1809

Qy  1771 TCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACT 1830
Db  1810 TTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGCGACCATT 1869

Qy  1831 ACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAG 1890
Db  1870 ACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCGGCTCTAG 1929

Qy  1891 GAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACC 1950
Db  1930 GAAGGAAAACTTGTTATGTACAATTTCCACAGAGATTTGATGGCATTGACTTGCACGATC 1989

Qy  1951 GATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTC 2010
Db  1990 GATATGCTAATCGGAACATAGTCTTCTTTGATATCAACATGAAAGGTCTGGATGGCATTC 2049

Qy  2011 AGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATG 2070
Db  2050 AGGGTCCAGTCTACGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGGATACG 2109

Qy  2071 ATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCT 2130
```

```
Db      2110 ATCC----------------------------TGTTTTGACTGAAGCTGATCTGGAGCCT 2141

Qy      2131 TTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA 2190

Db      2142 AACATTGTTGTTAAGAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATATGGAT 2201

Qy      2191 AGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTG 2250

Db      2202 AGTCAAAGCCGTA--TTATGAAGAGAACAGAATCTTCAGCTCCCATCTTCAACATGGAAG 2259

Qy      2251 AAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAA 2310

Db      2260 ACATCGAGGAGGGTATTGAAGGTTACGAGGATGAAAGGTCAGTGCTTATGTCCCAGAGGA 2319

Qy      2311 AATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATG 2370

Db      2320 AATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTCATTGCATCCACCTTTATGACTCAAG 2379

Qy      2371 GTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTA 2430

Db      2380 GTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTAAAGGAAGCTATCCATGTCATCA 2439

Qy      2431 GTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAG 2490

Db      2440 GTTGTGGATATGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGATCTATGGTTCAG 2499

Qy      2491 TTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACT 2550

Db      2500 TTACGGAGGATATTCTGACTGGGTTTAAANNNNNNNNNNAGGGGCTGGCAATCAATCTACT 2559

Qy      2551 GCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTC 2610

Db      2560 GCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATCGTCTTA 2619

Qy      2611 ACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTC 2670

Db      2620 ATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGCAGACATTGTCCTA 2679

Qy      2671 TTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCA 2730

Db      2680 TCTGGTATGGTTACAATGGACGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACCA 2739

Qy      2731 TCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTT 2790

Db      2740 TTGTTTATCCAATCACATCTGTTCCGCTTATCGCCTATTGTGTGCTTCCTGCTATCTGTC 2799

Qy      2791 TATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGT 2850

Db      2800 TTCTTACCAATAAATTTATCATTCCTGAGATTAGCAATTATGCTGGGATGTTCTTCATTC 2859

Qy      2851 CACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAA 2910

Db      2860 TTCTTTTCGCCTCCANNNNNGCCACTGGTATATTGGAGCTTAGATGGAGTGGTGTTGGCA 2919

Qy      2911 TTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCT 2970

Db      2920 TTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCTCTGCCCATCTCT 2979

Qy      2971 TTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGA 3030

Db      2980 TCGCAGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCAACTTCACAGTTA 3039

Qy      3031 CATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTA 3087

Db      3040 CCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTGGACCA 3099

Qy      3088 CCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCG 3147

Db      3100 GTTTGCTCATTCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAATGGTGGCAGGAA 3159

Qy      3148 TTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCT 3207

Db      3160 TTTCGTATGCCATTAACAGTGGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCTGTTCT 3219

Qy      3208 TTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAA 3267

Db      3220 TCTCGATCTGGGTGNTCCTCCATCTGTACACGTTCCTCAAGGGTCTCATGGGAAGGCAGA 3279

Qy      3268 ACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCC 3327

Db      3280 ACCGCACACCAACAATCGTCATTGTCTGGTCCATCCTTCTTGCATCTATCTTCTCCTTGC 3339

Qy      3328 TTTGGGTTCGGATTGATCCTTTCCT 3352

Db      3340 TGTGGGTGAAGATCGATCCTTTCAT 3364
```

RESULT 5
AY108113

LOCUS	AY108113 3763 bp mRNA linear HTC 16-OCT-2002
DEFINITION	Zea mays PCO126465 mRNA sequence.
ACCESSION	AY108113
VERSION	AY108113.1 GI:21211191
KEYWORDS	HTC.
SOURCE	Zea mays
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 3763)
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL	Unpublished (2002)
REFERENCE	2 (bases 1 to 3763)
AUTHORS	Coe,E.H.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES	Location/Qualifiers
source	1. .3763 /organism="Zea mays" /mol_type="mRNA" /db_xref="MaizeDB:637367" /db_xref="taxon:4577" /clone_lib="Maize Mapping Project/DuPont Cornsensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
BASE COUNT	863 a 984 c 1011 g 904 t 1 others
ORIGIN	

Query Match 32.8%; Score 1289.4; DB 11; Length 3763;
Best Local Similarity 64.7%; Pred. No. 1.2e-277;
Matches 2112; Conservative 1; Mismatches 1042; Indels 108; Gaps 9;

```
Qy       210 GACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGC 269

Db       344 GGCGACGCGGACGGCGTGAAGTCGGGGAGGCGCGGTGGCGGACAGGTGTGCCAGATCTGC 403

Qy       270 GGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCC 329

Db       404 GGCGACGGCGTGGGCACCACGGCGGAGGGGGACGTCTTCGCCGCCTGCGACGTCTGCGGG 463
```

```
Qy      330 TTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCC 389
Db      464 TTTCCGGTGTGCCGCCCCTGCTACGAGTACGAGCGCAAGGACGGCACGCAGGCGTGCCCC 523

Qy      390 CAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAG 449
Db      524 CAGTGCAAGACCAAGTACAAGCGCCACAAGGGGAGCCCGGCGATCCGTGGGGAGGAAGGA 583

Qy      450 GAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATG------ACTCG 503
Db      584 GACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATCTGGCAATGAGGACCAGAAG 643

Qy      504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db      644 CAGAAGATTGCCGACAGAATGCGCAGCTGGCGCATGAACGTTGGGGGCAGCGGGGATGTT 703

Qy      564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db      704 GGTCGCCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGTGGCGAGATT 763

Qy      624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGG 683
Db      764 CCTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTCAGGAGAAATCCCTGGTGCT 823

Qy      684 GGAAAGAGGATACATCCCCT---TCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGG 740
Db      824 TCCCCTGACCATCATATGATGTCCCCAACTGGGAACATTGGCAAGCGTGCTCCATTTCCC 883

Qy      741 TCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAG---------TGTTGCT 791
Db      884 TATGTGAACCATTCGCCAAATCCGTCAAGGGAGTTCTCTGGTAGCATTGGGAATGTTGCC 943

Qy      792 TGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA---------------------- 829
Db      944 TGGAAAGAGAGGGTTGATGGCTGGAAAATGAAGCAGGACAAGGGGACGATTCCCATGACG 1003

Qy      830 --------GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG-------- 873
Db     1004 AATGGCACAAGCATTGCTCCCTCTGAGGGTCGGGGTGTTGGTGATATTGATGCATCAACT 1063

Qy      874 -------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAA 926
Db     1064 GATTACAACATGGAAGATGCCTTATTGAACGACGAAACTCGACAGCCTCTATCTAGGAAA 1123

Qy      927 ATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTG 986
Db     1124 GTTCCACTTCCTTCCTCCAGGATAAATCCATACAGGATGGTCATTGTGCTGCGATTGATT 1183

Qy      987 GTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTG 1046
Db     1184 GTTCTAAGCATCTTCTTGCACTACCGTATCACAAATCCTGTGCGCAATGCATACCCATTA 1243

Qy     1047 TGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTC 1106
Db     1244 TGGCTTCTATCTGTTATATGTGAGATCTGGTTTGCTCTTTCGTGGATATTGGATCAGTTC 1303

Qy     1107 CCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGAC 1166
Db     1304 CCTAAGTGGTTTCCAATCAACCGGGAGACGTACCTTGATAGGCTGGCATTAAGGTATGAC 1363

Qy     1167 AAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCC 1226
Db     1364 CGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTCGTCAGTACAGTCGACCCA 1423

Qy     1227 TTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTAT 1286
Db     1424 ATGAAGGAGCCTCCTCTTGTCACTGCCAATACCGTGCTATCCATTCTTGCTGTGGATTAC 1483

Qy     1287 CCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAA 1346
Db     1484 CCTGTGGATAAGGTCTCTTGCTATGTATCTGATGATGGAGCTGCGATGCTGACATTTGAT 1543

Qy     1347 GCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAAT 1406

Db     1544 GCACTAGCTGAGACTTCAGAGTTTGCTAGAAAATGGGTACCATTTGTTAAGAAGTACAAC 1603

Qy     1407 ATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAG 1466
Db     1604 ATTGAACCTAGAGCTCCTGAATGGTACTTCTCCCAGAAAATTGATTACTTGAAGGACAAA 1663

Qy     1467 GTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAG 1526
Db     1664 GTGCACCCTTCATTTGTTAAAGACCGCCGGGCCATGAAGAGAGAATATGAAGAATTCAAA 1723

Qy     1527 GTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATG 1586
Db     1724 GTTAGGGTAAATGGCCTTGTTGCTAAGGCACAGAAAGTTCCTGAGGAAGGATGGATCATG 1783

Qy     1587 CAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTC 1646
Db     1784 CAAGATGGCACACCATGGCCAGGAAACAATACCMGGGACCATCCTGGAATGATTCAGGTT 1843

Qy     1647 TTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTAT 1706
Db     1844 TTCCTTGGTCACAGTGGTGGCCTTGATACTGAGGGCAATGAGCTACCCCGTTTGGTCTAT 1903

Qy     1707 GTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCA 1766
Db     1904 GTTTCTCGTGAAAAGCGTCCTGGATTCCAGCATCACAAGAAAGCTGGTGCCATGAATGCT 1963

Qy     1767 TTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGAT 1826
Db     1964 CTTGTTCGTGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAATCTTGATTGTGAT 2023

Qy     1827 CACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTA 1886
Db     2024 CACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCCTTATGGACCCTAAC 2083

Qy     1887 CTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCAT 1946
Db     2084 CTAGGAAGGAGTGTCTGCTACGTCCAGTTTCCCCAGAGATTCGATGGCATTGACAGGAAT 2143

Qy     1947 GACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGT 2006
Db     2144 GATCGATATGCCAACAGGAACACCGTGTTTTTCGATATTAACTTGAGAGGTCTTGATGGC 2203

Qy     2007 ATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGT 2066
Db     2204 ATCCAAGGACCAGTTTATGTCGGAACTGGCTGTGTTTTCAACCGAACAGCTCTATATGGT 2263

Qy     2067 TATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGG 2126
Db     2264 TATGAGCCCCCAATTAAGCAGAAG----------------------------AAGGGT 2293

Qy     2127 TGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACA 2186
Db     2294 GGTTTCTTGTCATCACTATGTGGCGGTAGG------AAGAAGGCAAGCAAATCAAAGAAG 2347

Qy     2187 GAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTT 2246
Db     2348 GGCTCGGACAAGAAGAAGTCGCAGAAGCATGTGGACAGTTCTGTGCCAGTATTCAACCTT 2407

Qy     2247 GGTGAAATTGACGAAGCTGCTCCAGGAGC------TGAGAATGAAAAGGCCGGTATTGTA 2300
Db     2408 GAAGATATAGAGGAGGGAGTTGAAGGCGCTGGATTTGACGACGAGAAATCACTTCTTATG 2467

Qy     2301 AATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTT 2360
Db     2468 TCTCAAATGAGCCTGGAGAAGAGATTTGGCCAGTCCGCAGCGTTTGTTGCCTCCACTCTG 2527

Qy     2361 CTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATA 2420
Db     2528 ATGGAGTATGGTGGTGTTCCTCAGTCCGCAACTCCGGAGTCTCTTCTGAAAGAAGCTATC 2587

Qy     2421 CATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATC 2480
```

```
Db       2588 CATGTTATAAGCTGTGGCTATGAGGACAAGACTGAATGGGGAACTGAGATCGGGTGGATC 2647

Qy       2481 TATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGG 2540
              || || || || |||||||| ||||| || || ||||||||||     | || ||||||
Db       2648 TACGGTTCTGTGACAGAAGACATTCTCACCGGATTCAAGATGCACGCGCGAGGCTGGCGG 2707

Qy       2541 TCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCA 2600
              || || |||||||| || || |||    || ||||| || ||||| ||   | ||||||||
Db       2708 TCGATCTACTGCATGCCCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTCG 2767

Qy       2601 GATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAAT 2660
              || |||||  |||||||||| |||||||||||||||||  | ||||||  ||||||||
Db       2768 GACCGTCTGAACCAGGTGCTCCGGTGGGCTCTTGGGTCCGTGGAGATCCTCTTCAGCCGG 2827

Qy       2661 CATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTAC 2720
              || ||||| || ||||| || || || ||  | || || ||  |||| |||||  | |||
Db       2828 CACTGCCCCCTGTGGTACGGCTACGGAGGGCGGCTCAAGTTCCTGGAGAGATTCGCGTAC 2887

Qy       2721 ATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCT 2780
              |||||| ||| | | || ||    || || || || ||  |    ||||| |   ||||
Db       2888 ATCAACACCACCATCTACCCGCTCACGTCCATCCCGCTTCTCATCTACTGCATCCTGCCC 2947

Qy       2781 GCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTG 2840
              ||||||||| |  | || || || || |||| ||||||| | |  ||  | |||||| |
Db       2948 GCCATCTGTCTGCTCACCGGAAAGTTCATCATTCCAGAGATCAGCAACTTCGCCAGCATC 3007

Qy       2841 TGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGT 2900
              |||||||| || || || |||| || || || ||| ||||||| || ||||| |||||
Db       3008 TGGTTCATCTCCCTCTTCATCTCGATCTTCGCCACGGGCATCCTGGAGATGAGGTGGAGC 3067

Qy       2901 GGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCC 2960
              || || || || || || |||||||||| |||||||||||||| || || ||  | |||
Db       3068 GGGGTGGGCATCGACGAGTGGTGGAGGAACGAGCAGTTCTGGGTGATCGGGGGCATCTCC 3127

Qy       2961 TCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGC 3020
              | |||||||| || |||||||||| || ||||||||  | || ||  | || || | |
Db       3128 GCGCACCTCTTCGCCGTGTTCCAGGGCCTGCTCAAGGTGCTGGCCGGCATCGACACCAAC 3187

Qy       3021 TTCACCGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTC 3077
              |||||||| || || ||||      |  || || |  || ||| | ||||| || |  |||
Db       3188 TTCACCGTCACCTCCAAGGCCTCGGACGAGGACGGCGACTTCGCGGAGCTGTACATGTTC 3247

Qy       3078 AAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTG 3137
              || ||||| ||  |  |||| || || |||||| | ||  |  | ||| |  | || ||
Db       3248 AAGTGGACGACGCTCCTGATCCCGCCCACCACCATCCTGATCATCAACCTGGTCGGCGTC 3307

Qy       3138 GTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGG 3197
              |||||  ||| | ||  | || ||||| | ||||||  ||||  |||||||| |||||  ||
Db       3308 GTCGCCGGCATCTCCTACGCCATCAACAGCGGATACCAGTCGTGGGGCCCGCTCTTCGGC 3367

Qy       3198 AAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTT 3257
              ||||| ||||| || || ||||| || ||||| || || || || ||||| ||  |  |
Db       3368 AAGCTCTTCTTCGCCTTCTGGGTCATCGTCCACCTGTACCCGTTCCTCAAGGGCCTCATG 3427

Qy       3258 GGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATC 3317
              || ||||| ||| | || || || || ||| |||||||| |||| |||||||| || |||
Db       3428 GGCAGGCAGAACCGCACCCCGACCATCGTCGTCGTCTGGGCCATCCTGCTGGCGTCCATC 3487

Qy       3318 TTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTT 3377
              |||||  | || |||||||| || || || |||    |    |   ||| |||  |
Db       3488 TTCTCCTTGCTGTGGGTTCGCATCGACCCCTTCACCACCCGCGTCACTGGCCCGGATACC 3547

Qy       3378 GAGGAGTGTGGTTTGGATTGCAA 3400
                ||  ||||||  |  | ||| |
Db       3548 CAGACGTGTGGCATCAACTGCTA 3570


RESULT 6
AY110079
LOCUS       AY110079                3696 bp    mRNA    linear   HTC 17-OCT-2002
DEFINITION  Zea mays CL1164_1 mRNA sequence.
ACCESSION   AY110079
VERSION     AY110079.1  GI:21214162
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3696)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 3696)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1. .3696
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:630055"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      831 a    766 c    880 g    945 t    274 others
ORIGIN

  Query Match             30.8%;  Score 1212.4;  DB 11;  Length 3696;
  Best Local Similarity   61.2%;  Pred. No. 2e-260;
  Matches 2004;  Conservative    0;  Mismatches 1162;  Indels  108;  Gaps    8;

Qy        210 GACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGC 269
              |||||||     ||   ||||  ||  |  | |   |  |||||||||||||||||| |||
Db        257 GACGGCGGCGACGCCACGAATTCGGGGAAGCATGTGGCCGGGCAGGTGTGCCAGATCTGC 316

Qy        270 GGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCC 329
              ||||||| ||| |||   ||  | | ||| |||| ||||   || ||| |||  |||| |
Db        317 GGCGACGGCGTGGGCACCGCGGCGGACGGCGACCTCTTCACCGCCTGCGACGTCTGCGGC 376

Qy        330 TTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCC 389
              |||||||| |||||    |||||||| |||||||||  ||| ||||| |||   |||||
Db        377 TTCCCCGTGTGCCGCCCATGCTACGAGTACGAGCGCAAGGACGGCACCCAGGCGTGCCCG 436

Qy        390 CAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAG 449
              |||||||||||||   |||||||||| ||| ||  |||  ||  ||   |||||| ||  |
Db        437 CAGTGCAAGACTAAGTACAAGCGCCACAAAGGGAGCCCACCAGTACACGGTGAGGAAAAT 496

Qy        450 GAGGACGGCGTCGATGACCTGGACAACGAGTTCAACT-----------GGGACGGCCAT 497
              ||||| |  |  | ||||   |  |  || | |||||           |  ||    ||
Db        497 GAGGATGTGGATGCTGACGATGTGAGTGACTACAACTACCAAGCATCTGGCAACCAGGAT 556

Qy        498 GACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT-------------- 543
              |  |||     | || |||   ||||||    | |  |  | ||
Db        557 CAGAAGCAAAAGATTGCTGAGAGAATGCTCACTTGGCGGACAAACTCACGTGGCAGTGAT 616

Qy        544 -------------------ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTC 584
                                  || || ||| |  ||  | |           | |
Db        617 ATTGGCCTGGCTAAGTATGACAGCGGTGAAATTGGGCATGGGAAGTATGACAGTGGTGAG 676
```

```
Qy       585 CAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCG 644
Db       677 ATCCCTCGTGGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCTGGA 736

Qy       645 GAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTT 704
Db       737 GCTTCCCCTGATCATATGATGTCTCCTGTTGGGAACATTGGCAGGCGTGGACATCAATTT 796

Qy       705 CCTTATGCGGAT-----------CCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCAT 753
Db       797 CCTTATGTAAATCATTCTCCAAACCCATCGAGGGAGTTCTCCGGTAGCCTTGGCAATGTT 856

Qy       754 CCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATT 813
Db       857 GCATGGAAAGAGAGGGTGGATGGATGGAAAATGAAGGATAAAGGTGCAATTCCTATGACC 916

Qy       814 GGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGA-------ATGATGGTGGTGGTGAT 866
Db       917 AATGGAACAAGCATTGCTCCATCAGAAGGGCGTGGAGTTGCTGATATTGATGCTTCTACT 976

Qy       867 GATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAA 926
Db       977 GATTATAACATGGAAGATGCCTTACTGAATGATGAAACTCGGCAACCTCTATCTAGAAAA 1036

Qy       927 ATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTG 986
Db      1037 GTGCCAATTCCTTCATCCAGAATAAATCCGTACAGAATGGTCATTGTGCTACGTTTGGCT 1096

Qy       987 GTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTG 1046
Db      1097 GTTCTATGCATATTCTTGCGCTACCGTATCACACATCCTGTGAACAATGCATATCCACTG 1156

Qy      1047 TGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTC 1106
Db      1157 TGGCTTTTATCCGTCATATGTGAGATCTGGTTTGCTTTGTCCTGGATTTTGGATCAGTTC 1216

Qy      1107 CCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGAC 1166
Db      1217 CCAAAGTGGTCCCCAATCAACCGTGAAACATACCTTGATAGACTGGCTTTAAGGTATGAC 1276

Qy      1167 AAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCC 1226
Db      1277 CGAGAAGGTGAACCATCTCAATTAGCTCCTGTTGATANNNNNGTCAGTACTGTGGATCCA 1336

Qy      1227 TTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTAT 1286
Db      1337 ATGAAGGAGCCTCCTCTTGTCACTGCAAATACTGTGCTTTCCATCCTTGCTGTCGATTAT 1396

Qy      1287 CCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAA 1346
Db      1397 CCGGTTGACAAGGTATCTTGCTATGTTTCGGATGATGGAGCTGCTATGCTGACTTTTGAT 1456

Qy      1347 GCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAAT 1406
Db      1457 GCTCTCTCTGAAACTTCAGAGTTTGCTAGAAAATGGGTTCCGTTCTGTAAGAAGTACAAC 1516

Qy      1407 ATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAG 1466
Db      1517 ATAGAGCCTANGGCCCCGGAATGGTACTTTGCTCAGAAAATTGATTACTTGAAAGACAAA 1576

Qy      1467 GTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAG 1526
Db      1577 GTTCAAACCTCATTTGTGAAAGAACGCCGGGNNNNNNNNNNNNNNNNNNNNNNNNNCAAA 1636

Qy      1527 GTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATG 1586
Db      1637 GTTCGTATCAATGGTCTTGTAGCCAAGGCACNNNNNGTTCCCGAGGAGGGATGGATCATG 1696

Qy      1587 CAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTC 1646
Db      1697 CAAGATGGTACACCTTGGCCTGGGAACAATACTAGGGACCATCCTGGAATGATTCAGGTT 1756

Qy      1647 TTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTAT 1706
```

```
Db      1757 TTCCTGGGTCACAGTGGAGGGCTTGACGTTGAAGGCAATGAACTTCCTCGTTTGGTTTAT 1816

Qy      1707 GTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCA 1766
Db      1817 GTGTCTCGTGNNNNNCGTCCTGGATTCCAACATCACAAGAAGGCTGGTGCCATGAATGCA 1876

Qy      1767 TTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGAT 1826
Db      1877 CTTGTTCGTGTATCAGCTGTCCTTACTAATGGGCAATACATGTTGAATCTTGATTGTGAC 1936

Qy      1827 CACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTA 1886
Db      1937 CACTACATCAATAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCCTTATGGACCCAAAC 1996

Qy      1887 CTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCAT 1946
Db      1997 CTAGGAAGGAATGTCTGTTATGTCCAATTTCCTCAGAGGTTTGATGGTATTGATAGGAAT 2056

Qy      1947 GACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGT 2006
Db      2057 GACCGATATGCAAACAGGAACACTGTGNNNNNCGATATTAACTTGAGAGGTCTTGACGGC 2116

Qy      2007 ATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGT 2066
Db      2117 ATTCAAGGGCCAGTTTATGTGGGAACTGGTTGTGTGTTTAACAGAACGGCCTTATATGGT 2176

Qy      2067 TATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGG 2126
Db      2177 TATGAGCCTCCAGTCAAGNNNNNNNNGCCA---------------------------- 2206

Qy      2127 TGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACA 2186
Db      2207 GGCTTCTTCTCTTCGCTTTGTNNNNNAAGGNNNNNGACGTCAAAATCTAAGAAGAGCTCG 2266

Qy      2187 GAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTT 2246
Db      2267 GAAAAGAAGAAGTCACA------TAGACACGCAGACAGTTCTGTACCAGTATTTAATCTC 2320

Qy      2247 GGTGAAATTGACGAAGCTGCTCCAGGAGC------TGAGAATGAAAAGGCCGGTATTGTA 2300
Db      2321 GAAGATATAGAGGAAGGGATTGAAGGTTCTCAGTTTGATGATGAGAAATCGCTGATTATG 2380

Qy      2301 AATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTT 2360
Db      2381 TCTCAAATGAGCTTGGAGAAGAGATTTGGCCAGTCCAGTGNNNNNGTAGCCTCTACTCTG 2440

Qy      2361 CTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATA 2420
Db      2441 ATGGAATATGGTGGTGTTCCACAATCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNTATT 2500

Qy      2421 CATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATC 2480
Db      2501 CATGTCATCAGCTGTGGCTATGAGGACAAAACTGACTGGGGAACTGAGATTGGGTGGATC 2560

Qy      2481 TATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGG 2540
Db      2561 TATGGTTCTGTTACAGAAGACATTCTCACCGGNNNNNNNNNNNNNNNNNNNNNNNNNNN 2620

Qy      2541 TCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCA 2600
Db      2621 NNNNNNNNNNNNNNNNCTAAGCGACCAGCTTTCAAGGGATCTGCTCCTATCAACCTTTCG 2680

Qy      2601 GATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAAT 2660
Db      2681 GATCGTTTGAATCAAGTGCTTCGGTGGGCTCTTGGTTCCATTGAAATTCTTTTCAGCAGG 2740

Qy      2661 CATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTAC 2720
Db      2741 CATTGTCCCATATGGTATGGCTATGGAGGCCGGCTTAAATTCCTGGAGAGATTTGCTTAT 2800

Qy      2721 ATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCT 2780
```

```
Db       2801 ATCAACACAACAATTTATCCACTCACATCAATCCCGCTCCTCCTGTACTGCATATTGCCA 2860

Qy       2781 GCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTG 2840

Db       2861 GCAGTTTGTCTTCTCACTGGGAAGTTCATCATCCCAAAGATTAGTAACCTAGAGAGTGTT 2920

Qy       2841 TGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGT 2900

Db       2921 TGGTTTATATCGCTCTTTATCTCAATCTTTGCCACTGGTATCCTTGNNNNNNNNNNNNNN 2980

Qy       2901 GGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCC 2960

Db       2981 NNNNNNNNCATTGATGAATGGTGGAGGAACGAGCAGTTCTGGGTCATTGGTGGTATTTCT 3040

Qy       2961 TCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGC 3020

Db       3041 GCGCATTTATTTGCCGTCTTCCAGGGTCTCCTGAAGGTGCTTGCTGGTATCGACACGAGC 3100

Qy       3021 TTCACCGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTC 3077

Db       3101 TTCACTGTCACCTCTAAGGCCACTGACGAAGAAGGTGATTTTGCCGAGCTCTACATGTTC 3160

Qy       3078 AAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTG 3137

Db       3161 AAGTGGACAACGCTTCTGATCCCACCAACCACTATTTTGATCATCAACCTGGTCGGCGTG 3220

Qy       3138 GTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGG 3197

Db       3221 GTCGCTGGCATTTCCTACGCAATCAATAGCGGTTACCAGTCATGGGGACCTCTTTTCGGG 3280

Qy       3198 AAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTT 3257

Db       3281 AAGCTCTTCTTTGCGTTCTGGGTGATTGTCCACCTGTACCCCTTCCTCAAGGGCCTCATG 3340

Qy       3258 GGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATC 3317

Db       3341 GGGAAGCAGAACCGCACGCCGACCATTGTCGTTGTCTGGGCTATCCTCCTTGCGTCGATC 3400

Qy       3318 TTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTT 3377

Db       3401 TTTTCCCTGATGTGGGTTCGTATCGATCCATTCACCACCCGGGTCACTGGTCCTGATATC 3460

Qy       3378 GAGGAGTGTGGTTTGGATTGCAACTAGGATGTCA 3411

Db       3461 GCGAAATGTGGCATCAACTGCTAGGATGAGCTGA 3494
```

RESULT 7
AY112236
LOCUS AY112236 3728 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1160_1 mRNA sequence.
ACCESSION AY112236
VERSION AY112236.1 GI:21216826
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3728)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3728)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES Location/Qualifiers
source 1. .3728
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
BASE COUNT 782 a 845 c 925 g 878 t 298 others
ORIGIN

Query Match 30.7%; Score 1209.6; DB 11; Length 3728;
Best Local Similarity 59.9%; Pred. No. 8.7e-260;
Matches 1963; Conservative 0; Mismatches 1210; Indels 102; Gaps 7;

```
Qy        210 GACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGC 269

Db        244 GGCGACGCGGACGGCGTGAAGNNNNNNNNNNNNNNNNNNNNNNNNNGTGCCAGATCTGC 303

Qy        270 GGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCC 329

Db        304 GGCGATGGCGTGGGCACTACGGCGGAGGGAGACGTCTTCACCGCCTGCGACGTCTGCGGG 363

Qy        330 TTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCC 389

Db        364 TTCCCGGTGTGCCGCCCCTGCTACGAGTACGAGCGCAAGGACGGCACACAAGCGTGNNNN 423

Qy        390 CAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAG 449

Db        424 NAGTGCNNNNNCAAGTACAAGCGCCACAAGGGGAGTCCAGCGATCCGAGGGGAGGAAGGA 483

Qy        450 GAGGACGGCGTCGATGACCTGGACAACGAGTTCAACT-----------GGGACGGCCAT 497

Db        484 GACGATACTGATGCCGATGATGCTAGCGACTTCAACTACCCTGCATCTGGCAATGACGAC 543

Qy        498 GACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGT 557

Db        544 CAGAAGCAGAAGATTGCTGACAGGATGCGCAGCTGGCGCATGAATGCTNNNNNCAGCGGG 603

Qy        558 GACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAAC 617

Db        604 GATGTTGGCCGCCCCAAGTATGACAGTGGTGAGATCGGGCTTACCAAGTACGACAGTGGT 663

Qy        618 GGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGT 677

Db        664 GAGATCCCTCGGGGATACATCCCNNNNNNNNNNNNNNNNNNNATTTCGGGAGAAATCCCT 723

Qy        678 GGTGGGGGAAAGAGGATACATCCCCT---TCCTTATGCGGATCCCAGCTTACCTGTGCAA 734

Db        724 GGTGCTTCCCCTGACCATCATATGATGTCTCCTACTGGGAACATTGGCAGGCGCGCCCCA 783

Qy        735 CCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAG---------T 785

Db        784 TTTCCCTATATGAATCATTCATCAAATCCGTCGAGGGAATTCTCTGGTAGCGTTGGGAAT 843

Qy        786 GTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA---------------- 829

Db        844 GTTGCCTGGAAAGAGAGGGTTGATGGCTGGAAAATGAAGCAGGACAAGGGAACAATTCCC 903

Qy        830 --------------GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG-- 873

Db        904 ATGACGAATGGCACAAGCATTGCTCCCTCTGAGGGCCGGGGTGTTGGTGATATTGATGCA 963
```

```
Qy    874 -------------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCC 920
Db    964 TCAACTGATTACAACATGGAAGATGCCTTATTAAACGATGAAACTCGCCAGCCTCTATCT 1023

Qy    921 AGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGG 980
Db   1024 AGGAAAGTTCCACTTCCTTCCTCCAGGATAAATCCATACAGGATGGTCATTGTGCTACGA 1083

Qy    981 CTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTT 1040
Db   1084 TTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAAATCCTGTGCGTAATGCATAC 1143

Qy   1041 GCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGAT 1100
Db   1144 CCACTGTGGCTTCTATCTGTTATATGTGAGATCTGGTTTGCTCTTTCCTGGATATTGGAT 1203

Qy   1101 CAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGG 1160
Db   1204 CAGTTTCCAAAGTGGTTTCCAATCAACCGCGAGACTTACCTTGATAGACTCGCATTAAGG 1263

Qy   1161 TTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTT 1220
Db   1264 TATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACANNNNNGTCAGTACTGTC 1323

Qy   1221 GATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTG 1280
Db   1324 GACCCAATGAAGGAGCCTCCTCTTGTCACTGCCAATACCGTGCTATCCATTCTCGCTGTG 1383

Qy   1281 GATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACG 1340
Db   1384 GACTATCCTGTGGATAAGGTCTCTTGCTATGTATCTGATGATGGAGCTGCTATGCTGACA 1443

Qy   1341 TTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGG 1400
Db   1444 TTTGATGCACTAGCTGAGACTTCAGAGTTTGCTAGAAAATGGGTGCCATTTGTTAAGAAG 1503

Qy   1401 TACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAA 1460
Db   1504 TACAACATTGAACCTAGAGCTCCTGAATGGTACTTCTCCCAGAAAATTGATTACTTGAAG 1563

Qy   1461 GACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAA 1520
Db   1564 GACAAAGTGCACCCTTCATTTGTTAAAGACCGCCGGGNNNNNNNNNNNNNNNNNNNNNNN 1623

Qy   1521 TTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGG 1580
Db   1624 NNCAAAATTAGGGTAAATGGCCTTGTTGCTAAGGCACNNNNNGTCCCTGAGGAAGGATGG 1683

Qy   1581 ACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATT 1640
Db   1684 ATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCAGGGACCATCCTGGAATGATT 1743

Qy   1641 CAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTG 1700
Db   1744 CAGGTTTTCCTTGGTCACAGTGGTGGTCTTGATACTGAGGGTAATGAGCTACCCCGTTTG 1803

Qy   1701 GTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATG 1760
Db   1804 GTCTATGTTTCTCGTGNNNNNCGTCCTGGATTCCAGCATCACAAGAAAGCTGGTGCCATG 1863

Qy   1761 AATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGAT 1820
Db   1864 AATGCTCTTGTCCGCGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAATCTTGAT 1923

Qy   1821 TGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGAC 1880
Db   1924 TGTGATCACTACATCAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCCTTATGGAT 1983

Qy   1881 CCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGAT 1940
Db   1984 CCTAACCTAGGAAGGAGTGTCTGCTATGTTCAGTTTCCCCAGAGGTTCGATGGTATTGAT 2043

Qy   1941 CGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTG 2000
```

```
Db   2044 AGGAATGATCGATATGCCAACAGGAACACCGTGNNNNNCGATATTAACTTGAGAGGTCTT 2103

Qy   2001 GATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTA 2060
Db   2104 GATGGCATCCAAGGACCAGTTTATGTGGGCACTGGCTGTGTTTTCAACAGAACAGCTCTA 2163

Qy   2061 TATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCC 2120
Db   2164 TATGGTTATGAGNNNNNAATTAAGCAAAAG---------------------------- 2193

Qy   2121 AAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCC 2180
Db   2194 AAGGGTGGTTTCTTGTCATCACTATGTGGNNNNNNNNNNNNNNNNNNNNNCAAAGAAG 2253

Qy   2181 AAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATAT 2240
Db   2254 GGCTCAGACAAGNNNNNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2313

Qy   2241 GCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTA 2300
Db   2314 NNNNNNNTAGAGGAGGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCACTTCTTATG 2373

Qy   2301 AATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTT 2360
Db   2374 TCTCAAATGAGCTTGGAGAAGAGATTTGGCCAATCTGCAGCTTTTGTTGCGTCCACTCTG 2433

Qy   2361 CTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATA 2420
Db   2434 ATGGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTATC 2493

Qy   2421 CATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATC 2480
Db   2494 CATGTCATAAGTTGTGGCTACGAGGACAAGATTGAATGGGGAACTGAGATTGGGTGGATC 2553

Qy   2481 TATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGG 2540
Db   2554 TATGGTTCTGTGACGGAAGATATTCTCACTGGGTTCAAGATGCACGCACGAGGCTGGCGG 2613

Qy   2541 TCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCA 2600
Db   2614 TCGATCTACTGCATGCCTAAGCGGCCGGCCTTCAAGGGATCGGCTCCCATCAATCTCTCA 2673

Qy   2601 GATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAAT 2660
Db   2674 GACCGTCTGAACCAGGTGCTCCGGTGGGCTCTCGGTTCAGTGGAAATCCTTTTCAGCCGG 2733

Qy   2661 CATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTAC 2720
Db   2734 CATTGNNNNNTATGGTACGGGTACGGAGGACGCCTGAAGTTCTTGGAGAGATTCGCCTAC 2793

Qy   2721 ATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCT 2780
Db   2794 ATCAACACCACCATCTACCCGCTCACGTCCCTCCCGCTCCTCATTTACTGTATCCTGCCT 2853

Qy   2781 GCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTG 2840
Db   2854 GCCATCTGCCTGCTCACGGGGAAGTTCATCATCCCAGAGATCAGCAACTTCGCTAGTATC 2913

Qy   2841 TGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGT 2900
Db   2914 TGGTTCATCTCTCTCTTCATCTCGATCTTCGCCACGGGTATCCTGGAGATGAGGTGGAGC 2973

Qy   2901 GGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCC 2960
Db   2974 GGCGTGGGCATCGACGAGTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCATCTCC 3033

Qy   2961 TCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGC 3020
Db   3034 GCCCACCTCTTCGCCGTCTTCCAGGGCCTCCTCAAGGTGCTTGCCGGCATCGACACCAAC 3093

Qy   3021 TTCACCGTGACATCAAAGGGTGGAGATGATGA---GGAGTTCTCAGAGCTATATACATTC 3077
```

```
Db       3094 TTCACCGTCACCTCCAAGGCCTCGGATGAAGACGGCGACTTCGCGGAGCTGTACATGTTC 3153

Qy       3078 AAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTG 3137
              || ||||| ||  |  |||| || || |||||| | ||  |  | ||| |  | || ||
Db       3154 AAGTGGACGACACTTCTGATCCCGCCCACCACCATCCTGATCATCAACCTGGTCGGCGTT 3213

Qy       3138 GTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGG 3197
              || || ||| | ||  | || ||||| | ||| ||  |||| ||||| || |||||  ||
Db       3214 GTTGCCGGCATCTCCTACGCCATCAACAGCGGGTACCAGTCGTGGGGTCCGCTCTTCGGC 3273

Qy       3198 AAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTT 3257
              ||||| ||||| || || |||||||| || || || || || ||  ||||| ||| |  |
Db       3274 AAGCTCTTCTTCGCCTTCTGGGTGATCGTTCACCTGTACCCGTTCCTCAAGGGTCTCATG 3333

Qy       3258 GGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATC 3317
              ||  |||| ||| | || || || || ||  | |||||| | || |||||||| || |||
Db       3334 GGTCGGCAGAACCGCACCCCGACCATCGTGGTTGTCTGGGCGATCCTGCTGGCGTCGATC 3393

Qy       3318 TTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTT 3377
              |||||  | || |||||||| || ||||| |||          |   ||| |||  |  |
Db       3394 TTCTCCTTGCTGTGGGTTCGCATCGATCCGTTCACCAACCGCGTCACTGGCCCGGATACT 3453

Qy       3378 GAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAG 3412
              ||||||  |  | |||  |   || ||
Db       3454 CGAACGTGTGGCATCAACTGCTAGGGAGGTGGAAG 3488
```

```
RESULT 8
AY104236
LOCUS       AY104236               2872 bp    mRNA    linear   HTC 16-OCT-2002
DEFINITION  Zea mays PCO121439 mRNA sequence.
ACCESSION   AY104236
VERSION     AY104236.1  GI:21207314
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 2872)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 2872)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1. .2872
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:637064"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      730 a    600 c    688 g    852 t      2 others
ORIGIN
```

```
Query Match             29.5%;  Score 1161.6;  DB 11;  Length 2872;
Best Local Similarity   68.4%;  Pred. No. 4.7e-249;
Matches 1679;  Conservative    1;  Mismatches  740;  Indels   34;  Gaps    4;

Qy        903 GCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGG 962
              ||| | | ||  ||      | | | ||||  |  |  ||| | || |||| |  ||| ||
Db         13 GCACGGCTACCTCTAAGTCGCATAGTTCCGATATCTCCAAACGAGCTTAACCTTTATCGG 72

Qy        963 ATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCAT 1022
              ||  | ||  || | |||||| |  |  |  | |||||||| || || ||  | |  |||
Db         73 ATCGTGATTGTTCTCCGGCTTATCATCCTATGTTTCTTCTTTCAATATCGTATAACTCAT 132

Qy       1023 CCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCC 1082
              || ||| | ||||| | ||  |||||||| |||||||||| |||||| | |||||||||
Db        133 CCAGTGGAAGATGCTTATGGGTTGTGGCTTGTATCTGTTATTTGTGAAGTTTGGTTTGCC 192

Qy       1083 ATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTA 1142
              |||||||| |||| ||||| |||||||||:||||  |||||  |  | || |||||||||
Db        193 TTGTCTTGGCTTCTAGATCAGTTCCCAAAGTGGTATCCTATCAACCGTGAAACTTACCTC 252

Qy       1143 GACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACT-TGCTCCAATTGA 1201
              ||  | ||  || |||| |  || | ||| ||  |||||||| ||  |  |||||||| ||
Db        253 GATAGACTTGCATTGAGATATGATAGGGAGGGTGAGCCATCCCAGTTGGGCTCCAATCGA 312

Qy       1202 TTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGT 1261
              | ||||||| ||||| || |||||  | ||||||||||||  |  | ||    || ||||||
Db        313 TGTCTTTGTTAGTACAGTGGATCCACTTAAGGAACCTCCTCTAATTACTGGCAACACTGT 372

Qy       1262 TCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGA 1321
              || || || ||| | |||||||| |||||||| || || || || ||||||||||||||
Db        373 CCTGTCCATTCTTGCTGTGGATTACCCTGTTGACAAAGTATCATGTTATGTTTCTGATGA 432

Qy       1322 TGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATG 1381
              ||| | || ||| | || |||||||| |||||||||| | || ||||||| ||||||
Db        433 CGGTTCAGCTATGTTGACTTTTGAAGCGCTATCTGAAACCGCAGAGTTTGCAAGGAAATG 492

Qy       1382 GGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACA 1441
              |||||| || |||||     |||||||||||||| | ||||||||||  |||||    |
Db        493 GGTTCCCTTTTGCAAGAAACACAATATTGAACCTAGGGCTCCAGAGTTTTACTTTGCTCG 552

Qy       1442 GAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAAT 1501
              |||||||| ||| | || |||||  |   | |    |||||  | || ||| | || ||
Db        553 AAAGATAGATTACCTAAAGGACAAAATACAACCTTCTTTTGTGAAAGAAAGGCGGGCTAT 612

Qy       1502 GAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAA 1561
              |||||| |||| ||| || ||||| ||   | ||| |||||  | |||||  ||||| || ||
Db        613 GAAGAGGGAGTGTGAAGAGTTCAAAGTACGGATCGATGCCCTTGTTGCAAAAGCGCAAAA 672

Qy       1562 AGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCG 1621
              | | ||||| || || ||||| |||   |||||  || ||:|||||||| || ||   | |
Db        673 AATACCTGAGGAGGGCTGGACCATGGCTGATGGCACTCCYTGGCCTGGGAATAACCCTAG 732

Qy       1622 TGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGG 1681
              |||||||| |||||||| || ||| | ||||| || || || ||||||   || ||
Db        733 AGATCATCCMGGAATGATCCAAGTATTCTTGGGCCACAGTGGTGGGCTTGACACGGATGG 792

Qy       1682 AAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCA 1741
              |||||  |||||||  | |||||||||||| | || ||  | |||||||   | || ||
Db        793 GAATGAGTTGCCACGGCTTGTTTATGTTTCTCGTGAAAAGAGGCCAGGCTTCCAGCACCA 852

Qy       1742 TAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCC 1801
              ||||| |||||||| |||||||| ||| | || || || ||||| || || |||| | |
Db        853 CAAGAAGGCTGGTGCCATGAATGCTTTGATTCGCGTATCAGCTGTCCTGACGAATGGTGC 912

Qy       1802 ATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGC 1861
              ||| |  | ||  |||||||||||||||| |||| | |||||| ||| | |  || ||
Db        913 TTATCTTCTTAATGTGGATTGTGATCACTACTTCAATAGCAGCAAAGCTCTTAGAGAGGC 972

Qy       1862 AATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCA 1921
              |||||||| |||||||| ||   |||||||| |||   |||||||| ||||| || ||
Db        973 TATGTGTTTCATGATGGATCCAGCACTAGGAAGGAAAACTTGCTATGTTCAGTTTCCACA 1032
```

```
Qy       1922 AAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGA 1981
Db       1033 AAGATTTGATGGTATAGACTTGCATGATCGATATGCAAACCGGAACATTGTCTTCTTTGA 1092

Qy       1982 TATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGT 2041
Db       1093 TATTAATATGAAGGGTCTAGATGGCATTCAAGGACCTGTTTATGTGGGAACAGGATGCTG 1152

Qy       2042 ATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAG 2101
Db       1153 TTTCAATAGGCAGGCCTTGTATGGCTATGAT----------------CCTGTATTGAC 1194

Qy       2102 GACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCA 2161
Db       1195 AGAAGCTGATTTGGAGCCTAACATTATCATTAAAAGTTGCTGTGGCGGAAGAAAAAAGAA 1254

Qy       2162 AAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGA 2221
Db       1255 GGACAAGAGCTATATTGATTCCAAAAACCGTGATA----------TGAAGAGAACAGA 1302

Qy       2222 GAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAA 2281
Db       1303 ATCTTCGGCTCCCATCTTCAACATGGAAGATATAGAAGAGGGATTTGAAGGTTACGAGGA 1362

Qy       2282 TGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGT 2341
Db       1363 TGAAAGGTCACTGCTTATGTCTCAGAAGAGCTTGGAGAAACGCTTTGGCCAGTCTCCAAT 1422

Qy       2342 TTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTC 2401
Db       1423 TTTTATTGCATCCACCTTTATGACTCAAGGTGGCATACCCCCTTCAACAAACCCAGGTTC 1482

Qy       2402 TCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGG 2461
Db       1483 CCTGCTAAAGGAAGCTATACATGTCATTAGTTGTGGATATGAGGATAAAACAGAATGGGG 1542

Qy       2462 AAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGAT 2521
Db       1543 GAAAGAGATCGGATGGATATATGGCTCTGTTACTGAAGATATTTTAACTGGTTTCAAGAT 1602

Qy       2522 GCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTC 2581
Db       1603 GCATGCAAGAGGTTGGATATCCATCTACTGCATGCCACTTCGGCCTTGCTTCAAGGGTTC 1662

Qy       2582 TGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTAT 2641
Db       1663 TGCTCCGATTAATCTTTCTGATCGTCTCAACCAAGTGTTACGCTGGGCTCTTGGTTCAGT 1722

Qy       2642 TGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATT 2701
Db       1723 TGAAATTCTACTTAGCAGACACTGTCCTATCTGGTATGGTTACAATGGAAGGCTAAAGCT 1782

Qy       2702 TTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTT 2761
Db       1783 TCTGGAGAGACTGGCATACATCAACACCATTGTTTATCCAATTACATCTATCCCACTAGT 1842

Qy       2762 GGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCT 2821
Db       1843 AGCATACTGCGTCCTTCCTGCTATCTGTTTACTCACCAACAAATTTATTATTCCTGCGAT 1902

Qy       2822 GAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCAT 2881
Db       1903 TAGCAATTATGCTGGGGCGTTCTTCATCCTGCTTTTTGCTTCCATCTTCGCCACTGGTAT 1962

Qy       2882 CCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTG 2941
Db       1963 TTTGGAGCTTCGATGGAGTGGTGTTGGCATTGAGGATTGGTGGAGAAATGAGCAGTTTTG 2022

Qy       2942 GGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCAT 3001
Db       2023 GGTCATTGGTGGCACCTCTGCACATCTCTTTGCTGTGTTCCAAGGTCTCTTAAAAGTGCT 2082

Qy       3002 AGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGG---GTGGAGATGATGAGGAGTT 3058
Db       2083 AGCAGGGATCGACACAAACTTCACGGTCACATCAAAGGCAACCGATGATGATGGTGATTT 2142

Qy       3059 CTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCT 3118
Db       2143 TGCTGAGCTGTATGTGTTCAAGTGGACAACTCTTCTGATCCCCCCCACCACTGTGCTTGT 2202

Qy       3119 ATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTC 3178
Db       2203 GATTAACCTGGTTGGTATAGTGGCTGGAGTGTCGTATGCTATCAACAGTGGCTACCAATC 2262

Qy       3179 ATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCC 3238
Db       2263 ATGGGGTCCACTATTCGGGAAGCTGTTCTTTGCAATCTGGGTGATCCTCCACCTCTACCC 2322

Qy       3239 CTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTC 3298
Db       2323 TTTCCTGAAGGGTCTCATGGGGAAGCAGAACCGCACACCGACCATCGTCATTGTTTGGTC 2382

Qy       3299 CATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352
Db       2383 CGTCCTTCTTGCTTCCATATTCTCGCTGCTGTGGGTGAAGATCGACCCCTTCAT 2436
```

RESULT 9
CC379931/c

```
LOCUS       CC379931                 877 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHOP50TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa498J04,
            genomic survey sequence.
ACCESSION   CC379931
VERSION     CC379931.1  GI:30855570
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
            ,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa498J04"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
BASE COUNT      290 a    220 c    174 g    193 t
ORIGIN
```

Query Match 20.6%; Score 811; DB 29; Length 877;
Best Local Similarity 99.9%; Pred. No. 1.3e-170;
Matches 822; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       2818 AGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGA 2877
Db        822 AGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCA-TTTTGCTACGA 764

Qy       2878 GCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGT 2937
```

```
Db        763 GCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGT 704

Qy       2938 TCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGG 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 TCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGG 644

Qy       2998 TCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGT 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         643 TCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGT 584

Qy       3058 TCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTC 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         583 TCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTC 524

Qy       3118 TATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGT 3177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         523 TATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGT 464

Qy       3178 CATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATC 3237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         463 CATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATC 404

Qy       3238 CCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGT 3297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         403 CCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGT 344

Qy       3298 CCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGA 3357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         343 CCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGA 284

Qy       3358 AGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCAT 3417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         283 AGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCAT 224

Qy       3418 CAGCTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAG 3477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         223 CAGCTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAG 164

Qy       3478 TGTCTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACT 3537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         163 TGTCTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACT 104

Qy       3538 TAATATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGG 3597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         103 TAATATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGG 44

Qy       3598 CCGTGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTC 3640
             |||||||||||||||||||||||||||||||||||||||||||
Db          43 CCGTGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTC 1
```

RESULT 10
CD441181/c

```
LOCUS       CD441181                758 bp    mRNA    linear   EST 03-JUN-2003
DEFINITION  EL01N0553C04.g Endosperm_5 Zea mays cDNA, mRNA sequence.
ACCESSION   CD441181
VERSION     CD441181.1  GI:31356824
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
            Messing,J.
  TITLE     Sequencing of the maize endosperm ESTs
  JOURNAL   Unpublished
COMMENT     Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm_5"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      236 a    177 c    144 g    201 t
ORIGIN
```

Query Match 18.6%; Score 731.4; DB 14; Length 758;
Best Local Similarity 99.2%; Pred. No. 8.5e-153;
Matches 735; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       3144 GGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTA 3203
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 GGGGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTA 699

Qy       3204 TTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGG 3263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         698 TTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGG 639

Qy       3264 CAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCG 3323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 CAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCG 579

Qy       3324 CTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAG 3383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         578 CTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAG 519

Qy       3384 TGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGCTCCCCCAATCTGCATATGCTTG 3443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         518 TGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGCTCCCCCAATCTGCATATGCTTG 459

Qy       3444 AAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGTCTGTAGATAAGAGACATGAAATG 3503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         458 AAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGTCTGTAGATAAGAGACATGAAATG 399

Qy       3504 TCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAATATCTGAGAGATATACTGGGGGA 3563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         398 TCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAATATCTGAGAGATATACTGGGGGA 339

Qy       3564 AAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCGTGGAATACAGCATATGCAAGTGT 3623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 AAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCGTGGAATACAGCATATGCAAGTGT 279

Qy       3624 TTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATATAGATGGGCTGAGCCGAACAGCA 3683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         278 TTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATATAGATGGGCTGAGCCGAACAGCA 219

Qy       3684 AGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACTTGGTTCTCAATAAGGCAGGCAGG 3743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 AGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACTTGGTTCTCAATAAGGCAGGCAGG 159

Qy       3744 AATGCATCTGCCAGTGGAACAGAGCAACCTGCACATTATTTATGTATGCCTGTTCATTGG 3803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         158 AATGCATCTGCCAGTGGAACAGAGCAACCTGCACATTATTTATGTATGCCTGTTCATTGG 99

Qy       3804 AGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAAAACAGAATATTAGCATTAATCTA 3863
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||| |||||
Db          98 AGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAAACCAGAATATTAGCATTGATCTA 39

Qy       3864 TAGTTAATTAAAGTATGTAAA 3884
             ||||||||||||||||   |||
Db          38 TAGTTAATTAAAGTAAAAAAA 18
```

```
RESULT 11
CB633016
LOCUS       CB633016                856 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEb11J21.f OSIIEb Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEb11J21 5', mRNA sequence.
ACCESSION   CB633016
VERSION     CB633016.1  GI:29628005
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 856)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 11  row: J  column: 21
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .856
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEb11J21"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEb"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (PO6-6-3)"
BASE COUNT      222 a    165 c    203 g    266 t
ORIGIN

  Query Match             17.9%;  Score 704.6;  DB 14;  Length 856;
  Best Local Similarity   89.0%;  Pred. No. 8.7e-147;
  Matches  761;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy       2219 AGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGA 2278
             ||| || |||||||||||||||| |||||||||||| || | ||||||||||||||||
Db          2 AGAAAATCAATCCCCTGCATATGCACTTGGTGAAATTGAGGAGGGTGCTCCAGGAGCTGA 61

Qy       2279 GAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTC 2338
             |||||||||| |||||||||||||||||| ||  | ||||||||||||||||||||||
Db         62 AAATGAAAAGGCTGGTATTGTAAATCAACAGAAGCTGGAAAAGAAATTTGGCCAATCTTC 121

Qy       2339 TGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGC 2398
             ||||||||| |||||||||||| |||||||||||||||||||| ||||| ||||| ||
Db        122 TGTTTTTGTCGCATCCACACTTCTTGAGAATGGTGGAACCTTGAAAAGTGCGAGTCCAGC 181

Qy       2399 TTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTG 2458
             |||||| | || ||||| || |||||||| || ||||| || ||||| ||||||||||
Db        182 ATCTCTTCTAAAGGAAGCAATTCATGTCATAAGCTGTGGGTACGAAGATAAGACAGACTG 241

Qy       2459 GGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAA 2518
             ||| |||||||||||||||||| |||||||||||||||||||||||||||||||| ||
Db        242 GGGGAAAGAGATTGGCTGGATCTACGGATCAGTTACAGAAGATATTCTAACTGGTTTTAA 301

Qy       2519 GATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGG 2578
             |||||| || |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        302 GATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAACTACCTGCATTCAAAGG 361

Qy       2579 TTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTC 2638
             ||||||||| || |||||||| |||||||||||||| || ||||||||||||||||||||
Db        362 TTCTGCACCACTCAATCTTTCTGATCGTCTTCACCAAGTTCTTCGGTGGGCTCTTGGGTC 421

Qy       2639 TATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAA 2698
             | |||| || |||||||||| |||||||||||||||||||| |||||||| || || ||
Db        422 TGTTGAAATATTCTTCAGCAACCATTGCCCTCTTTGGTATGGATATGGTGGTGGACTAAA 481

Qy       2699 ATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCT 2758
             || |||||||||||||||||||||| || |||||||||||||||| ||||| ||||| ||
Db        482 ATGTTTGGAAAGATTTTCCTACATTAATTCCATCGTGTATCCTTTTACATCCATTCCACT 541

Qy       2759 CTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGA 2818
             |||||||||||||||||||| |||||||||| | || ||||| ||||||||||||||||
Db        542 CTTGGCTTACTGTACATTGCCAGCCATCTGTTTGCTAACTGGGAAGTTTATCACTCCAGA 601

Qy       2819 GCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAG 2878
             ||| | |||||||||||| || |||||||||||||||||||||||||||||||| |
Db        602 GCTTACGAATGTTGCCAGTCTCTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACTGG 661

Qy       2879 CATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTT 2938
             |||||| |||||||||||||||||||||| |||||||| |||||||| ||||| || ||
Db        662 CATCCTGGAAATGAGATGGAGTGGTGTTGGCATTGATGATTGGTGGAGAAATGAACAATT 721

Qy       2939 CTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGT 2998
             |||||||||||||| || || || |||||||||||| |||||||||||| | |||||
Db        722 TTGGGTCATTGGAGGGGTCTCTTCGCACCTCTTTGCTTTGTTCCAGGGACTTTTGAAGGT 781

Qy       2999 CATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTT 3058
             |||||||||| | |||||| || ||||| |||||||||| ||||| || ||||| |||||
Db        782 CATAGCTGGTATAGATACGAGTTTCACTGTGACATCAAAAGGTGGGGACGATGAAGAGTT 841

Qy       3059 CTCAGAGCTATATAC 3073
              ||||||||||||||
Db        842 CTCAGAGCTATATAC 856


RESULT 12
CB645001
LOCUS       CB645001                859 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSJNEb06H07.f OSJNEb Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEb06H07 5', mRNA sequence.
ACCESSION   CB645001
VERSION     CB645001.1  GI:29639992
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 06  row: H  column: 07
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
```

```
     source          1. .859
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEb06H07"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEb"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (Che 86061
                     )"
BASE COUNT      238 a    172 c    199 g    250 t
ORIGIN

  Query Match             17.4%;  Score 686.2;  DB 14;  Length 859;
  Best Local Similarity   88.2%;  Pred. No. 1.2e-142;
  Matches  758;  Conservative    0;  Mismatches   98;  Indels    3;  Gaps    1;

Qy       2062 ATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCA 2121
              |||||||||||||||||||| ||||||||  ||  ||  |||||  |||||||||||||| |
Db          1 ATGGTTATGATGCCCCCAAAACGAAGAAGCCTCCGTCGAGGACCTGCAACTGCTGGCCTA 60

Qy       2122 AGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAG---CAAAAGAAGACTACCAAAC 2178
              | ||||||  ||||||||||||||||||||  | ||||||    | |||||  |||||  ||||
Db         61 AATGGTGCATTTGCTGTTGCTGCTTTGGTGACAGGAAGAGCAAGAAGAAAACTACGAAAC 120

Qy       2179 CCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCAT 2238
              | ||||||||||||||| ||    ||  || |||||||  || |||  || ||||||||||||||
Db        121 CTAAAACAGAGAAGAAGAAAAGATCTTTCTTCAAGAGAGCAGAAAATCAATCCCCTGCAT 180

Qy       2239 ATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTG 2298
              ||||  ||||||||||||||| || | |||||||||||||||  ||||||||||  |||||||
Db        181 ATGCACTTGGTGAAATTGAGGAGGGTGCTCCAGGAGCTGAAAATGAAAAGGCTGGTATTG 240

Qy       2299 TAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACAC 2358
              |||||||||| ||   | |||||||||||||||||||||||||||||||||  |||||||||
Db        241 TAAATCAACAGAAGCTGGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTCGCATCCACAC 300

Qy       2359 TTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTA 2418
              ||||  |||||||||||||||||||| |||||  |||||  ||  ||||||  |  ||  ||||| |
Db        301 TTCTTGAGAATGGTGGAACCTTGAAAAGTGCGAGTCCAGCATCTCTTCTAAAGGAAGCAA 360

Qy       2419 TACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGA 2478
              | ||||||||  ||  |||||  ||  |||||  ||||||||||||||  ||||||||||||||||
Db        361 TTCATGTCATAAGCTGTGGGTACGAAGATAAGACAGACTGGGGGAAAGAGATTGGCTGGA 420

Qy       2479 TCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGC 2538
              ||||  |||||||||||||||||||||||||||||||  ||||||||  ||  ||||||||||
Db        421 TCTACGGATCAGTTACAGAAGATATTCTAACTGGTTTTAAGATGCACTGCCATGGTTGGC 480

Qy       2539 GGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTT 2598
              ||||||||||||||||||||||||    ||||||||||||||||||||  ||  ||||||||
Db        481 GGTCAATTTACTGCATACCTAAACTACCTGCATTCAAAGGTTCTGCACCACTCAATCTTT 540

Qy       2599 CAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCA 2658
              | ||||||||||||||  ||  ||||||||||||||||||||  ||||  ||  ||||||||||
Db        541 CTGATCGTCTTCACCAAGTTCTTCGGTGGGCTCTTGGGTCTGTTGAAATATTCTTCAGCA 600

Qy       2659 ATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCT 2718
              | ||||||||||||||||||||  ||||||||  ||  ||  ||||  ||||||||||||||||||
Db        601 ACCATTGCCCTCTTTGGTATGGATATGGTGGTGGACTAAAATGTTTGGAAAGATTTTCCT 660

Qy       2719 ACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGC 2778
              ||||  ||  ||||||||||||||||  |||||  |||||  ||||||||||||||||||||||
Db        661 ACATTAATTCCATCGTGTATCCTTTTACATCCATTCCACTCTTGGCTTACTGTACATTGC 720

Qy       2779 CTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCC 2838
              | ||||||||||  | ||  |||||  ||||||||||||||||  |  ||||||||||  |
Db        721 CAGCCATCTGTTTGCTAACTGGGAAGTTTATCACTCCAGAGCTTACGAATGTTGCCAGTC 780

Qy       2839 TGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGA 2898
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACTGGCATCCTGGAAATGAGATGGA 840

Qy       2899 GTGGTGTTGGAATTGATGA 2917
              ||||||||||  ||||||||
Db        841 GTGGTGTTGGCATTGATGA 859


RESULT 13
CB627056
LOCUS       CB627056                 864 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEb01L20.f OSIIEb Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEb01L20 5', mRNA sequence.
ACCESSION   CB627056
VERSION     CB627056.1  GI:29622045
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 864)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 01  row: L  column: 20
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .864
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEb01L20"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEb"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (PO6-6-3)"
BASE COUNT      259 a    166 c    211 g    228 t
ORIGIN

  Query Match             17.4%;  Score 683.2;  DB 14;  Length 864;
  Best Local Similarity   87.7%;  Pred. No. 5.4e-142;
  Matches  758;  Conservative    0;  Mismatches  103;  Indels    3;  Gaps    1;

Qy       1744 AGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCAT 1803
              ||||  ||||||||  ||  ||  ||||||||||||||  ||||||||||||||  ||||||||||
Db          1 AGAAGGCTGGTGCCATCAACGCATTGGTCCGAGTATCTGCTGTACTAACTAATGCTCCAT 60

Qy       1804 ATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAA 1863
              |  ||  |  ||||||||||||||||  ||||||||||  |||||||||||||||||||||||
Db         61 ACATGCTCAACTTGGATTGTGATCATTACATCAACAATAGCAAGGCTATAAAGGAAGCAA 120

Qy       1864 TGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAA 1923
              ||||||||||||||||||||||  |  ||||||||||||||  |||||  ||  ||||||||||
Db        121 TGTGTTTTATGATGGACCCTTTGGTTGGAAAGAAGGTTTGTTATGTGCAATTCCCTCAAA 180

Qy       1924 GATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATA 1983
```

```
Db        181 GGTTTGATGGGATTGATCGTCATGATCGATACGCCAACAGGAACGTGGTCTTCTTCGATA 240

Qy       1984 TCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTAT 2043

Db        241 TCAACATGAAAGGCTTGGATGGTATCCAAGGCCCCATCTATGTTGGTACTGGATGTGTCT 300

Qy       2044 TTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGA 2103

Db        301 TTAGAAGGCAGGCATTGTATGGTTATGATGCCCCCAAAACGAAGAAGCCTCCGTCGAGGA 360

Qy       2104 CTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAG---C 2160

Db        361 CCTGCAACTGCTGGCCTAAATGGTGCATTTGCTGTTGCTGCTTTGGTGACAGGAAGAGCA 420

Qy       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220

Db        421 AGAAGAAAACTACGAAACCTAAAACAGAGAAGAAGAAAAGATCTTTCTTCAAGAGAGCAG 480

Qy       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280

Db        481 AAAATCAATCCCCTGCATATGCACTTGGTGAAATTGAGGAGGGTGCTCCAGGAGCTGAAA 540

Qy       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340

Db        541 ATGAAAAGGCTGGTATTGTAAATCAACAGAAGCTGGAAAAGAAATTTGGCCAATCTTCTG 600

Qy       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400

Db        601 TTTTTGTCGCATCCACACTTCTTGAGAATGGTGGAACCTTGAAAAGTGCGAGTCCAGCAT 660

Qy       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460

Db        661 CTCTTCTAAAGGAAGCAATTCATGTCATAAGCTGTGGGTACGAAGATAAGACAGACTGGG 720

Qy       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520

Db        721 GGAAAGAGATTGGCTGGATCTACGGATCAGTTACAGAAGATATTCTAACTGGTTTTAAGA 780

Qy       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580

Db        781 TGCACTGCCATGGGTGGCGGTCAATTTACTGCATACCTAAACTACCTGCATTCAAAGGTT 840

Qy       2581 CTGCACCTCTGAATCTTTCAGATC 2604

Db        841 CTGCACCACTCAATCTTTCTGATC 864
```

```
RESULT 14
CD437450
LOCUS       CD437450                862 bp    mRNA    linear   EST 03-JUN-2003
DEFINITION  EL01N0501B03.b Endosperm_5 Zea mays cDNA, mRNA sequence.
ACCESSION   CD437450
VERSION     CD437450.1  GI:31353093
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 862)
  AUTHORS   Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
            Messing,J.
  TITLE     Sequencing of the maize endosperm ESTs
  JOURNAL   Unpublished
COMMENT     Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T3.
```

```
FEATURES             Location/Qualifiers
     source          1. .862
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm_5"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      172 a    216 c    218 g    256 t
ORIGIN
```

Query Match 17.1%; Score 672.4; DB 14; Length 862;
Best Local Similarity 87.4%; Pred. No. 1.4e-139;
Matches 736; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```
Qy       2586 CCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAG 2645

Db         11 CCTCTCAATCTTTCCGATCGTCTTCACCAGGTTCTTCTTTGGGCTCTTGGTTCAATTGAA 70

Qy       2646 ATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTG 2705

Db         71 ATTTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGTGGTGGACTAAAGTTCCTG 130

Qy       2706 GAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCT 2765

Db        131 GAAAGGTTTTCGTACATTAACTCCATCGTATACCCTTGGACATCTATCCCGCTCTTGGCC 190

Qy       2766 TACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAAT 2825

Db        191 TATTGCACATTGCCTGCCATCTGCTTGCTGACAGGGAAATTTATCACGCCAGAGCTTAAC 250

Qy       2826 AATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTA 2885

Db        251 AATGTTGCCAGCCTGTGGTTCATGTCACTTTTCATCTGCATTTTTGCTACGAGCATCCTG 310

Qy       2886 GAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTC 2945

Db        311 GAAATGAGATGGAGTGGTGTAGGCATCGATGACTGGTGGAGAAACGAGCAGTTTTGGGTC 370

Qy       2946 ATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCT 3005

Db        371 ATTGGAGGCGTGTCTTCACATCTCTTTGCTGTGTTCCAGGGACTCCTCAAGGTCATAGCT 430

Qy       3006 GGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAG 3065

Db        431 GGTGTAGACACGAGCTTCACTGTGACATCCAAGGGTGGAGACGACGAGGAGTTCTCAGAG 490

Qy       3066 CTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAAC 3125

Db        491 CTGTACACATTCAAATGGACGACCCTTCTGATACCTCCGACAACCCTGCTCCTACTGAAC 550

Qy       3126 TTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGC 3185

Db        551 TTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAACGGATATGAATCATGGGGC 610

Qy       3186 CCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTC 3245

Db        611 CCCCTGTTCGGGAAGCTCTTCTTTGCATTTTGGGTGATCGTCCATCTTTACCCGTTCCTC 670

Qy       3246 AAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTG 3305

Db        671 AAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTGTCATTGTCTGGTCCATCCTC 730

Qy       3306 CTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGAT 3365

Db        731 CTGGCTTCGATCTTCTCGCTGCTTTGGGTCCGGATCGACCCTTTCCTTGCGAAGGATGAT 790

Qy       3366 GGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAGCTCCC 3425

Db        791 GGTCCCCTGTTTGGAGAGTGTGGTCTGGATTGCAACTAAGAGGTCAGCACGTGGACTTCC 850
```

```
Qy       3426 CC 3427
              ||
Db        851 CC 852


RESULT 15
CC365622/c
LOCUS       CC365622                900 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  PUHBR52TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa414J08,
            genomic survey sequence.
ACCESSION   CC365622
VERSION     CC365622.1  GI:30835022
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
            ,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa414J08"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
BASE COUNT      270 a    213 c    198 g    219 t
ORIGIN

  Query Match             17.0%;  Score 670.6;  DB 29;  Length 900;
  Best Local Similarity   88.1%;  Pred. No. 3.6e-139;
  Matches  808;  Conservative    0;  Mismatches   79;  Indels   30;  Gaps    6;

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
              ||||  |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        900 TGTTTCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACTGTGACATCAA 841

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
              |||||||||||||||| || |||||||||| || ||||| ||||||||||||||||||||
Db        840 AGGGTGGAGATGATGACGAATTCTCAGAGCTGTACACATTTAAATGGACTACCTTATTGA 781

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
              |||||||||||||| ||||| |||||||||||||||||||||||||| ||||| ||||
Db        780 TACCTCCTACCACTTTGCTCCTATTGAACTTCATTGGTGTGGTCGCTGGTGTTTCCAATG 721

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTT-TGGGAAGCTATTCTTTGCATTT 3215
              | ||||||||| ||||| ||||||||||||||||||  |||||||| ||||||||||||
Db        720 CAATCAATAATGGATACGAGTCATGGGGCCCCCTCTTCGGGGAAGCTCTTCTTTGCATTT 661

Qy       3216 TGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACA 3275
              |||||||||||||||| |||||||| |||||||||||||||| ||||||||||||||||
Db        660 TGGGTGATTGTCCATCTGTATCCCTTCCTCAAAGGTTTGGTTGGGAGGCAAAACAGGACA 601

Qy       3276 CCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTT 3335
              || || |||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db        600 CCGACAATTGTCATCGTCTGGTCCATCCTGCTGGCTTCAATCTTCTCACTCCTTTGGGTT 541

Qy       3336 CGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGAT 3395
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CGGATTGACCCTTTCCTTGCAAAGGATGATGGTCCACTTCTTGAGGAGTGTGGTTTGGAT 481

Qy       3396 TGCAACTA--GGATGTCAGTGCATCAGCTCCCCCAATCTGCATATGCTTGAAGTATATTT 3453
              ||||||||  |||||||||  |||||||||||||||| || |||| ||||||||| |||
Db        480 TGCAACTAGGGGATGTCAGCACATCAGCTCCCCCAATCAGCGTATGATTGAAGTATTTTT 421

Qy       3454 TCTGGTGTTTGTCCCCATATTCAGTG-------TCTGTAGATAAGAGACATGAAATGTCC 3506
              | || ||||||||||||||| |  |       ||||||| ||||||||| ||||||||
Db        420 GCCGGCGTTTGTCCCCATATTTACAGTGCCCCCTCTGTAGGTAAGAGACAAGAAATGTCT 361

Qy       3507 CAAGTTTCTTTTGATCCATGGTGAACCTACTTAATATCTGAGAGATATACTGG-GGGAAA 3565
              |  ||| |||||||||||||||||||||        || |  | |||| ||||| ||||||
Db        360 CCCGTTCCTTTTGATCCATGGTGAACCTCGACCTAATATCTGGGATACACTGGAGGGAAA 301

Qy       3566 ATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCGTGGAATACAGCATATGCAAGTGTTT 3625
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        300 ATGGAGGCTGCGGCAATCCTTGTGCAGATGGGCCGTGGAATACAGCATATGCAAGTGTTT 241

Qy       3626 GATTGTGCAGCATTCTTTATTACTTGGTCGCAATATAGATGGGCTGAGCCGAACAGCAAG 3685
              |||||||||||||||||||||||||||| |||||  |  ||| | ||||||||| |||||||
Db        240 GATTGTGCAGCATTCTTTATTACTTGGGCGCAAAACTGATCGTCTGAGCCGAGCAGCAAG 181

Qy       3686 GTATTTTGATTCTGCACTGCTCCCGTGTACAAACTTGGTTCTCAATAAGGCAGGCAGGAA 3745
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GTATTTTGATCCTGCACTGCTCCCGTGTACAAACTTGGTTCTCAATAAGGCAGGCAGGAA 121

Qy       3746 --TGCATCTGCCAGTGGAACAGAGCAACCTGCACATTATTTATGTATGCCTGTTCATTGG 3803
                |||||||||||||||||||||| |||||||||  ||||||||||
Db        120 TGTGCATCTGCCAGTGGAACAGAGGAACCTGCACAGTATTTATGTA-------------- 75

Qy       3804 AGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAAAACAGAATATTAGCATTAATCTA 3863
              ||||||||||||||||||||||||||||||||||||||||||||||| | |  ||||||
Db         74 ---CAAACTTCATTACATGTTCGTCTATACTAGGAAAAACAGAATATTTGTACCAATCTA 18

Qy       3864 TAGTTAATTAAAGTATG 3880
              || || ||||||||||
Db         17 TACTTGATTAAAGTATG 1


Search completed: December 16, 2003, 00:49:16
Job time : 7651 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 14:57:33 ; Search time 958 Seconds
(without alignments)
11090.802 Million cell updates/sec

Title: US-09-720-383C-9
Perfect score: 3936
Sequence: 1 cttctccctcgtcggtgcgg.....ctgagttggttttgtgaaaa 3936

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3936	100.0	3936	21	AAZ58266	Corn cellulose syn
2	3936	100.0	3969	21	AAZ99503	DNA encoding a mai
3	3936	100.0	3969	21	AAZ99518	DNA encoding a mai
4	2592.6	65.9	3776	21	AAZ58263	Corn cellulose syn
5	2546.8	64.7	3568	21	AAZ99491	DNA encoding a mai
6	2241	56.9	3786	21	AAZ58265	Corn cellulose syn
7	2241	56.9	3813	21	AAZ99509	DNA encoding a mai
8	2241	56.9	3813	21	AAZ99524	DNA encoding a mai
9	1509	38.3	3255	24	ABZ14725	Arabidopsis thalia
10	1489.8	37.9	3444	22	AAC83798	Arabidopsis thalia
11	1431.8	36.4	3255	24	ABZ12754	Arabidopsis thalia
12	1429.4	36.3	3828	19	AAV06566	Arabidopsis cellul
13	1411.4	35.9	3851	21	AAA67114	Pinus radiata cell
14	1378.4	35.0	3725	21	AAZ99500	DNA encoding a mai
15	1378.4	35.0	3725	21	AAZ99506	DNA encoding a mai
16	1378.4	35.0	3725	21	AAZ99521	DNA encoding a mai
17	1364	34.7	3753	21	AAZ99515	DNA encoding a mai
18	1364	34.7	3753	21	AAZ99530	DNA encoding a mai
19	1364	34.7	3780	21	AAZ99497	DNA encoding a mai
20	1331.8	33.8	2890	21	AAZ58270	Soybean cellulose
21	1327.6	33.7	3673	19	AAV06568	Arabidopsis cellul
22	1326.6	33.7	3603	19	AAV06565	Arabidopsis cellul
23	1321.4	33.6	3704	21	AAZ99533	DNA encoding a mai
24	1289.4	32.8	3746	21	AAZ99512	DNA encoding a mai
25	1289.4	32.8	3746	21	AAZ99527	DNA encoding a mai
26	1289.4	32.8	3773	21	AAZ99494	DNA encoding a mai
27	1279.4	32.5	3799	24	AAS16458	Corn cDNA encoding
28	1272.4	32.3	3198	21	AAC49550	Arabidopsis thalia
29	1269.2	32.2	3614	19	AAV06567	Arabidopsis cellul
30	1254.2	31.9	3517	21	AAZ58268	Soybean cellulose
31	1246.2	31.7	3311	19	AAV08373	Cellulose synthase
32	1170	29.7	2830	24	AAS16455	Corn cDNA encoding
33	1137.6	28.9	3207	19	AAV08372	Cellulose synthase
34	1136.4	28.9	3328	19	AAV34432	Cotton cellulose s
35	1118.8	28.4	3747	21	AAA67145	Eucalyptus grandis
36	1106.2	28.1	3232	22	AAC65448	Populus tremuloide
37	998.8	25.4	2306	19	AAV34433	Cotton cellulose s
38	884.6	22.5	2248	19	AAV06562	Arabidopsis cellul
39	863	21.9	2033	19	AAV08381	Cellulose synthase
40	859.6	21.8	2125	21	AAZ58269	Soybean cellulose
41	856.4	21.8	2055	21	AAA67111	Eucalyptus grandis
42	749.2	19.0	1860	21	AAA67081	Eucalyptus grandis
43	741.4	18.8	1029	21	AAZ58272	Wheat cellulose sy
44	701.8	17.8	1734	21	AAZ58271	Wheat cellulose sy
45	664.2	16.9	1741	19	AAV06569	Rice cellulose syn

ALIGNMENTS

RESULT 1
AAZ58266
ID AAZ58266 standard; cDNA; 3936 BP.
XX
AC AAZ58266;
XX
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA clone p0122.ckamh70rc.
XX
KW Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 144..3404
FT /*tag= a
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15871.
XX
PR 14-JUL-1998; 98US-0092844.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX

DR WPI; 2000-182431/16.
DR P-PSDB; AAY58835.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
XX
PS Claim 9; Page 49-50; 93pp; English.
XX
CC The present sequence is that of full-length cDNA clone
CC p0122.ckamh70rc encoding corn cellulose synthase (CS, see AAY58835).
CC The cDNA clone was isolated from a cDNA library produced from corn
CC pith tissue collected from internode subtending ear node 5 days
CC after pollination. The invention relates to isolated nucleic acid
CC fragments encoding plant CS and to CS polypeptides. It also relates
CC to the construction of a chimeric gene encoding all or a portion of
CC the CS, in sense or antisense orientation, where expression of the
CC gene results in altered levels of the CS in transformed host cells.
CC The host cells can be used to screen compounds for their ability to
CC inhibit CS activity. CS nucleic acids are also useful for producing
CC transgenic plants having altered levels of CS, and hence altered
CC levels of fibre. CS may also serve as a target for the development
CC of novel herbicides.
XX
SQ Sequence 3936 BP; 989 A; 838 C; 1018 G; 1091 T; 0 other;

Query Match 100.0%; Score 3936; DB 21; Length 3936;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480

Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540

Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600

Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660

Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720

Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780

Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840

Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900

Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960

Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020

Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080

Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140

Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200

Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260

Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320

Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380

Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440

Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500

Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560

Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620

Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680

Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
```

```
Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800

Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860

Qy       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920

Qy       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980

Qy       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040

Qy       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100

Qy       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160

Qy       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220

Qy       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280

Qy       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340

Qy       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400

Qy       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460

Qy       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520

Qy       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580

Qy       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640

Qy       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700

Qy       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760

Qy       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820

Qy       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880

Qy       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940

Qy       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000

Qy       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060

Qy       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120

Qy       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180

Qy       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240

Qy       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300

Qy       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360

Qy       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420

Qy       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480

Qy       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540

Qy       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600

Qy       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660

Qy       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720

Qy       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780

Qy       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840

Qy       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900

Qy       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             |||||||||||||||||||||||||||||||||||||
```

```
Db       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936


RESULT 2
AAZ99503
ID   AAZ99503 standard; DNA; 3969 BP.
XX
AC   AAZ99503;
XX
DT   03-JUL-2000  (first entry)
XX
DE   DNA encoding a maize cellulose synthase.
XX
KW   Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW   transgenic plant; plant breeding marker; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             144..3398
FT                   /*tag=  a
FT                   /product= "cellulose synthase"
FT                   /note= "no termination codon given"
XX
PN   WO200009706-A2.
XX
PD   24-FEB-2000.
XX
PF   16-AUG-1999;   99WO-US18760.
XX
PR   17-AUG-1998;   98US-0096822.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Dhugga KS,  Helentjaris TG,  Bowen BA,  Wang X;
XX
DR   WPI; 2000-224343/19.
DR   P-PSDB; AAY84111.
XX
PT   New genes which encode maize cellulose synthase polypeptides in plants
PT   useful for modulating the expression of cellulose synthase in plants
PT   and to produce transgenic plants expressing the novel protein  -
XX
PS   Claim 1; Page 113-118; 119pp; English.
XX
CC   The present sequence encodes a maize cellulose synthase polypeptide.
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 other;

  Query Match             100.0%;  Score 3936;  DB 21;  Length 3969;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3936;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480

Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540

Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600

Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660

Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720

Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780

Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840

Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900

Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960

Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020

Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080

Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
```

```
Db      1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Qy      1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Db      1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Qy      1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Db      1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Qy      1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Db      1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Qy      1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Db      1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Qy      1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Db      1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Qy      1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Db      1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Qy      1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Db      1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Qy      1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Db      1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Qy      1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Db      1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Qy      1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Db      1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Qy      1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Db      1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Qy      1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Db      1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Qy      1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db      1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Qy      1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db      1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Qy      1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db      1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Qy      2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db      2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Qy      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Qy      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Qy      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Qy      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Db      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Qy      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Db      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Qy      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Db      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Qy      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Db      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Qy      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Db      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Qy      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Db      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Qy      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Db      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Qy      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Db      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Qy      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Db      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Qy      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Db      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Qy      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Db      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Qy      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Db      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Qy      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Db      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Qy      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Db      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Qy      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Db      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Qy      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Db      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Qy      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Db      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
```

```
Qy      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360

Qy      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420

Qy      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480

Qy      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540

Qy      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600

Qy      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660

Qy      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720

Qy      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780

Qy      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840

Qy      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900

Qy      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             ||||||||||||||||||||||||||||||||||||
Db      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 3
AAZ99518
ID AAZ99518 standard; DNA; 3969 BP.
XX
AC AAZ99518;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key   Location/Qualifiers
FT CDS   144..3398
FT   /*tag= a
FT   /product= "cellulose synthase"
FT   /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84116.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 1; Page 152-157; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 other;

Query Match 100.0%; Score 3936; DB 21; Length 3969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy        61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy       121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy       181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy       241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy       301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy       361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy       421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Db       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Db       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Db       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Db       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Db       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Db       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Db       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Db       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
```

```
Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Db       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Db       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Db       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Db       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Db       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Qy       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Qy       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Qy       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Qy       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Qy       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Qy       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Qy       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Qy       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Db       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Qy       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Db       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Qy       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Db       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Qy       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Db       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Qy       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Db       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Qy       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Db       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
```

```
Qy       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700

Qy       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760

Qy       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820

Qy       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880

Qy       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940

Qy       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000

Qy       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060

Qy       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120

Qy       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180

Qy       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240

Qy       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300

Qy       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360

Qy       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420

Qy       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480

Qy       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540

Qy       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600

Qy       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660

Qy       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720

Qy       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780

Qy       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840

Qy       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900

Qy       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             ||||||||||||||||||||||||||||||||||||
Db       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 4
AAZ58263
ID AAZ58263 standard; cDNA; 3776 BP.
XX
AC AAZ58263;
XX
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA.
XX
KW Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 2..3448
FT /*tag= a
FT /partial
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15871.
XX
PR 14-JUL-1998; 98US-0092844.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
DR WPI; 2000-182431/16.
DR P-PSDB; AAY58832.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
XX
PS Claim 9; Page 37-38; 93pp; English.
XX
CC The present sequence is that of a contig of cDNA clones and encodes
CC a portion (see AAY58832) of corn cellulose synthase (CS). The cDNA
CC clones were isolated from corn cob, developing tassel, seedling,
CC shoot culture and leaf tissue cDNA libraries on the basis of
CC homology to Arabidopsis and cotton CS sequences. The invention
CC relates to isolated nucleic acid fragments encoding plant CS and to
CC CS polypeptides. It also relates to the construction of a chimeric
CC gene encoding all or a portion of the CS, in sense or antisense
CC orientation, where expression of the gene results in altered levels
CC of the CS in transformed host cells. The host cells can be used to
CC screen compounds for their ability to inhibit CS activity. CS
CC nucleic acids are also useful for producing transgenic plants having
CC altered levels of CS, and hence altered levels of fibre. CS may
CC also serve as a target for the development of novel herbicides.
XX
SQ Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 other;

Query Match 65.9%; Score 2592.6; DB 21; Length 3776;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2885; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

```
Qy       104 CTCTGCCTTCGTCGGCGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCT 163
Db       139 CGCTTGCCCCGGGATCTGGAGATCTGGTAGCGCCAGGGGGATGGAGGCCAGCGCCGGGCT 198

Qy       164 GGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGG 223
Db       199 GGTCGCCGGCTCGCACAACCGGAACGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGG 258

Qy       224 GCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGG 283
Db       259 GCCGAAGCCCATGGACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGG 318

Qy       284 CCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCG 343
Db       319 GCGCAACCCCGACGGGGAGCCGTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCG 378

Qy       344 GGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCG 403
Db       379 GGACTGCTACGAGTACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCG 438

Qy       404 ATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGA 463
Db       439 CTTCAAGCGCCTCAAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGA 498

Qy       464 TGACCTGGACAACGAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTC 520
Db       499 CGACCTGGAGAACGAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTC 558

Qy       521 CATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGC 580
Db       559 CATGCTCCACGCCCACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCC 618

Qy       581 TTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCC 640
Db       619 ATTCCACCCCATCCCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCC 678

Qy       641 ACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGGAAAGAGGATACATCC 700
Db       679 GCCGGACCAGCACGCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGAAGAGGATTCACCC 738

Qy       701 CCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGA 760
Db       739 TCTCCCGTACGCGGATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGA 798

Qy       761 TCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCA 820
Db       799 TCTCGCCGCATATGGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCA 858

Qy       821 GAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGC 880
Db       859 GAAGCAGGAGAGGATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGC 918

Qy       881 TGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATC 940
Db       919 AGATCTACCACTAATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTC 978

Qy       941 AAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTT 1000
Db       979 AAGCCAAATCAACCCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTT 1038

Qy      1001 CTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGT 1060
Db      1039 CTTCCACTACCGAGTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGT 1098

Qy      1061 TATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCC 1120
Db      1099 GATCTGTGAAATTTGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCC 1158

Qy      1121 TATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCC 1180
Db      1159 TATCGAGAGGGAAACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCC 1218

Qy      1181 ATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCC 1240
Db      1219 TTCTCAACTCGCCCCTGTTGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCC 1278

Qy      1241 TTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGT 1300
Db      1279 ATTGGTCACTGCTAATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGT 1338

Qy      1301 TTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAAC 1360
Db      1339 TTCATGCTACGTTTCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAAC 1398

Qy      1361 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGC 1420
Db      1399 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGC 1458

Qy      1421 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTT 1480
Db      1459 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTT 1518

Qy      1481 TGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGC 1540
Db      1519 TGTTAGAGAACGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGC 1578

Qy      1541 CTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCC 1600
Db      1579 CTTGGTTGCTAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCC 1638

Qy      1601 CTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAG 1660
Db      1639 ATGGCCCGGAAATAATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAG 1698

Qy      1661 CGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAA 1720
Db      1699 TGGTGGCCATGATGTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAA 1758

Qy      1721 ACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1780
Db      1759 ACGGCCAGGCTACAACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1818

Qy      1781 TGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAA 1840
Db      1819 TGCTGTACTAACTAATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAA 1878

Qy      1841 CAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGT 1900
Db      1879 TAGTAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAAGAAAGT 1938

Qy      1901 TTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAA 1960
Db      1939 TTGCTATGTGCAGTTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAA 1998

Qy      1961 CCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAAT 2020
Db      1999 CAGAAATGTTGTCTTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAAT 2058

Qy      2021 TTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAA 2080
Db      2059 TTATGTGGGTACTGGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAA 2118

Qy      2081 AACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTG 2140
Db      2119 AACAAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTG 2178

Qy      2141 CTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAA 2194
Db      2179 CTGTTTTGGTAACAGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAA 2238

Qy      2195 AAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAAT 2254
```

```
Db       2239 GATAAAGAAACTTTTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAAT 2298

Qy       2255 TGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATT 2314

Db       2299 TGATGAAGCCGCTCCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTT 2358

Qy       2315 AGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGG 2374

Db       2359 GGAAAAGAAATTTGGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGG 2418

Qy       2375 AACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTG 2434

Db       2419 AACCCTGAAGAGTGCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGTTG 2478

Qy       2435 TGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTAC 2494

Db       2479 TGGATATGAAGACAAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCAC 2538

Qy       2495 AGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCAT 2554

Db       2539 AGAAGATATTCTTACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCAT 2598

Qy       2555 ACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCA 2614

Db       2599 ACCTAAACGGGCCGCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTCTTCACCA 2658

Qy       2615 GGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTG 2674

Db       2659 GGTTCTTCGGTGGGCTCTTGGTTCAATTGAAATTTTCTTCAGCAACCACTGCCCTCTCTG 2718

Qy       2675 GTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGT 2734

Db       2719 GTATGGGTATGGTGGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGT 2778

Qy       2735 GTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATT 2794

Db       2779 ATACCCTTGGACATCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCT 2838

Qy       2795 GACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACT 2854

Db       2839 GACAGGGAAATTTATCACGCCAGAGCTTAACAATGTTGCCAGCCTCTGGTTCATGTCACT 2898

Qy       2855 TTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGA 2914

Db       2899 TTTCATCTGCATTTTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGA 2958

Qy       2915 TGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGC 2974

Db       2959 TGACTGGTGGAGAAACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGC 3018

Qy       2975 TGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATC 3034

Db       3019 TGTGTTCCAGGGACTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATC 3078

Qy       3035 AAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATT 3094

Db       3079 CAAGGGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCT 3138

Qy       3095 GATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAA 3154

Db       3139 GATACCTCCGACAACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAA 3198

Qy       3155 TGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATT 3214

Db       3199 TGCGATCAACAACGGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATT 3258

Qy       3215 TTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGAC 3274

Db       3259 TTGGGTGATCGTCCATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGAC 3318

Qy       3275 ACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGT 3334

Db       3319 GCCAACGATTGTCATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGT 3378

Qy       3335 TCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGA 3394

Db       3379 CCGGATCGACCCGTTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGA 3438

Qy       3395 TTGCAACTAGGATGTCAGTGCATCAGCTCCCCC 3427

Db       3439 TTGCAACTAGGAGGTCAGCACGTGGACTTCCCC 3471
```

```
RESULT 5
AAZ99491
ID   AAZ99491 standard; DNA; 3568 BP.
XX
AC   AAZ99491;
XX
DT   03-JUL-2000  (first entry)
XX
DE   DNA encoding a maize cellulose synthase.
XX
KW   Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW   transgenic plant; plant breeding marker; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             63..3236
FT                   /*tag=  a
FT                   /product= "cellulose synthase"
FT                   /note= "no termination codon given"
XX
PN   WO200009706-A2.
XX
PD   24-FEB-2000.
XX
PF   16-AUG-1999;   99WO-US18760.
XX
PR   17-AUG-1998;   98US-0096822.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Dhugga KS,  Helentjaris TG,  Bowen BA,  Wang X;
XX
DR   WPI; 2000-224343/19.
DR   P-PSDB; AAY84107.
XX
PT   New genes which encode maize cellulose synthase polypeptides in plants
PT   useful for modulating the expression of cellulose synthase in plants
PT   and to produce transgenic plants expressing the novel protein  -
XX
PS   Claim 1; Page 82-87; 119pp; English.
XX
CC   The present sequence encodes a maize cellulose synthase polypeptide.
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
```

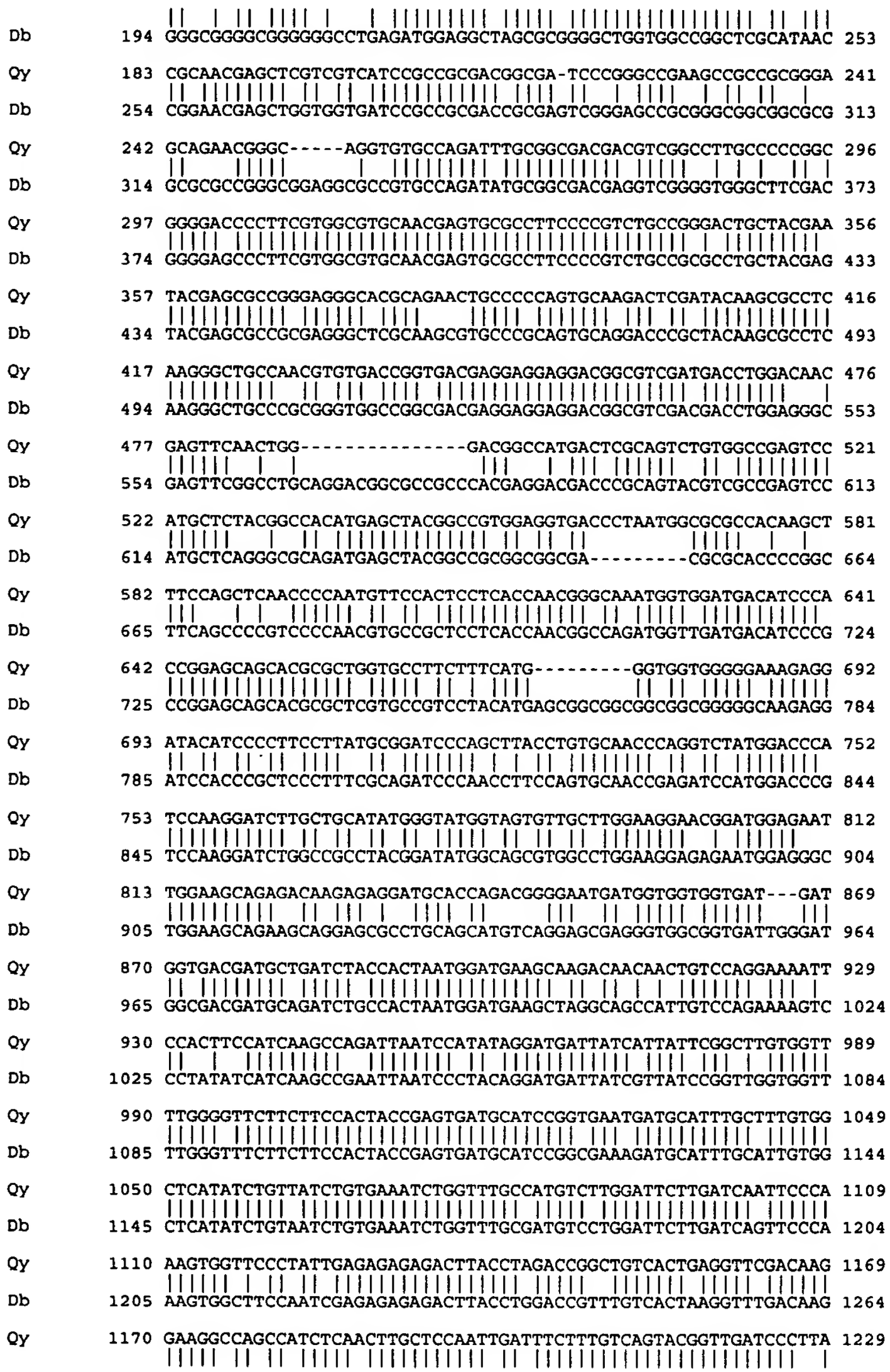
```
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence 3568 BP; 879 A; 825 C; 941 G; 922 T; 1 other;

  Query Match            64.7%;  Score 2546.8;  DB 21;  Length 3568;
  Best Local Similarity  86.8%;  Pred. No. 0;
  Matches 2829;  Conservative    0;  Mismatches  422;  Indels    9;  Gaps    2;

Qy       177 CACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCG 236
             | ||  ||    |||||||||||||||||||||||| ||||| || ||||||||||    |
Db         6 CCCACGCGTCCGGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGGGCCGAAGCCCATG 65

Qy       237 CGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
             |||  |||||| |||||||||||||||||||||||||||||| || |    ||||| |
Db        66 GACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGGGCGCAACCCCGAC 125

Qy       297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
             |||||  || |||||||| |||||||||||||||||||||| ||||||||||||||||||
Db       126 GGGGAGCCTTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCGGGACTGCTACGAG 185

Qy       357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
             |||||||||||| |||||||||||||||||||||||||||||||| || | ||||||| ||
Db       186 TACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCGCTTCAAGCGCTTC 245

Qy       417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
             |||||  |||    || ||| |||| |||||||||||||||||||||| |||||||| |||
Db       246 AAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGACGACCTGGAGAAC 305

Qy       477 GAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGC 533
             ||||||||||||   |||    || |||||  ||||   | ||||||||||||||| ||| |
Db       306 GAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTCCATGCTCCACGCC 365

Qy       534 CACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAAC 593
             |||||||||||||||| || |  ||||   | |||| ||| ||  | ||||| | || |
Db       366 CACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCCATTCCACCCCATC 425

Qy       594 CCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCAC 653
             |||||||||||  ||||||||||||  || |||||  |||||||||||| |||||||||||
Db       426 CCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCCGCCGGACCAGCAC 485

Qy       654 GCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCG 713
             || || |||||  || ||| ||||||| |||||  |||||||| || ||  || ||  |||
Db       486 GCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGAAGAGGATTCACCCTCTCCCGTACGCG 545

Qy       714 GATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATAT 773
             |||||||  |  | ||||||||||| |||||||||||| |||||||||| || ||||||
Db       546 GATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGATCTCGCCGCATAT 605

Qy       774 GGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGG 833
             || || || || || || ||||||||  |||||||||  ||||||||||  || ||||||
Db       606 GGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCAGAAGCAGGAGAGG 665

Qy       834 ATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTA 893
             ||||||||||||  |||| |||||  || ||  |||||||||||  |||||  ||||||||||||
Db       666 ATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGCAGATCTACCACTA 725

Qy       894 ATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAAT 953
             ||||||||||| |||||  |  | |||||||  ||  ||  ||  |||||  ||||||||  ||  ||
Db       726 ATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATCAAC 785

Qy       954 CCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGA 1013
             ||  |||||||||||||  ||  ||||||||  |||||||||  |  ||||||||||||||||||
Db       786 CCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTTCTTCCACTACCGA 845

Qy      1014 GTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATC 1073
             ||||||||||||||  ||||||||||||||||  ||||||||||||||  ||||||||||
Db       846 GTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGTGATCTGTGAAATT 905

Qy      1074 TGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAG 1133
             |||||||||||||||||||||||||| || ||  ||||||||||  |||||  |||||  ||
Db       906 TGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCCTATCGAGAGGGAA 965

Qy      1134 ACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCT 1193
             ||  |||  ||  |||||||||  |||||||||  |||||  |||||||||||  ||||||
Db       966 ACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCCTTCTCAACTCGCC 1025

Qy      1194 CCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACA 1253
             ||  |||||||||||||||||||||||||||||||||||||| ||||||||||  |||||||  |
Db      1026 CCTGTTGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCCATTGGTCACTGCT 1085

Qy      1254 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTT 1313
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||  |||||  |||
Db      1086 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGTTTCATGCTACGTT 1145

Qy      1314 TCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCA 1373
             |||||||||||||||||||  |||||  ||  ||||||||||  |||||||||||||||||||||
Db      1146 TCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAACATCTGAATTTGCA 1205

Qy      1374 AAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTAC 1433
             ||||||||||||||||||||||| | ||  |  ||||  |||||  ||||||||||||||||
Db      1206 AAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGCTCCAGAGTGGTAC 1265

Qy      1434 TTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGG 1493
             ||||||||||||||||||||  ||||||||||||||||  ||||||||||||||  ||  ||
Db      1266 TTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTTTGTTAGAGAACGG 1325

Qy      1494 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAA 1553
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||
Db      1326 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGCCTTGGTTGCTAAA 1385

Qy      1554 GCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAAC 1613
             ||||  ||  ||||||||  ||||||||||||||||  ||||||||  ||  |||||  |||||
Db      1386 GCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCCATGGCCCGGAAAT 1445

Qy      1614 AATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGAC 1673
             |||||  |||||||||||||||||||||||||||||  ||||||||  |||||  ||  |||| |||
Db      1446 AATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAGTGGTGGCCATGAT 1505

Qy      1674 TGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTAT 1733
                ||  |||||||| |||||  |||||||||||||||||  |||||  |||||  ||||||||
Db      1506 GTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAAACGGCCAGGCTAC 1565

Qy      1734 AACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACA 1793
             ||||||||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db      1566 AACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACT 1625

Qy      1794 AATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATA 1853
             ||||||||  ||||||  | ||||||||||||||||||||  |||||  ||  ||  |||||||||
Db      1626 AATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAATAGTAAGGCTATA 1685

Qy      1854 AAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAG 1913
             ||||||||||||||||||||||||||||||||||||  |||||  ||  |||||||| ||||||||||  |||
Db      1686 AAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAAGAAAGTTTGCTATGTGCAG 1745

Qy      1914 TTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTC 1973
             ||  |||||||||||||||||||||||||||||||||||||||  ||||||||||||  | |||||||||
Db      1746 TTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAACAGAAATGTTGTC 1805

Qy      1974 TTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACT 2033
             |||||  |||||||||||||||||||||||||||||||||  |||||  ||||||||||||  ||||||
Db      1806 TTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAATTTATGTGGGTACT 1865

Qy      2034 GGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCA 2093
             ||||||||  || ||||||||||||||||||||||  ||  |||||  ||||||||||||||||||
Db      1866 GGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAAAACAAAGAAGCCA 1925

Qy      2094 CCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAAT 2153
             ||||||||  ||||||||||||||||  |||||||||  ||||||||||||||  |||||  ||
Db      1926 CCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTTGGTAAC 1985

Qy      2154 AGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAAAAAGTTATTATTT 2207
             ||||||   |||||||  |||||  ||   | |   |      ||||  |       | ||
Db      1986 AGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAAGATAAAGAAACTT 2045

Qy      2208 TTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCT 2267
             ||  ||||||  |  ||  ||  |||  ||||||||||||||||||||||||||||  |||||  |||
```

```
Db       2046 TTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAATTGATGAAGCCGCT 2105

Qy       2268 CCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTT 2327
             |||||||||||| |||||||||||  ||||||||||||||||| || || ||||||||||||
Db       2106 CCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTTGGAAAAGAAATTT 2165

Qy       2328 GGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGT 2387
             |||||  |||||  ||||||||| |||||||||||| |||||||||||||| ||||||||
Db       2166 GGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGGAACCCTGAAGAGT 2225

Qy       2388 GCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGAC 2447
             ||  |||||  |||||||||  ||||  |||||||||||||||||  ||||||||  |||||||||
Db       2226 GCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGTTGTGGATATGAAGAC 2285

Qy       2448 AAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTA 2507
             ||  ||||  ||||||||||||  |||||  |||||  ||||||||||  ||||||||||||||
Db       2286 AAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCACAGAAGATATTCTT 2345

Qy       2508 ACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTT 2567
             |||||  ||  |||||||||  ||  ||||||||||||||||||||||||||||||||||||||
Db       2346 ACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGCC 2405

Qy       2568 GCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGG 2627
             ||  ||||||||||| ||||||||  ||||||||  ||||||  ||||||||||  |||||||||
Db       2406 GCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTTTTCACCAGGTTCTTCGGTGG 2465

Qy       2628 GCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGT 2687
             ||||||||  ||  |||||  ||  ||  ||||||||  ||  ||||||||  ||||||||||||||||
Db       2466 GCTCTTGGTTCAATTGAAATTTTGTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGT 2525

Qy       2688 GGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACA 2747
             ||  ||  ||  ||  ||   |||||||  |||||  |||||  ||||||||||  ||  |||||||||
Db       2526 GGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGTATACCCTTGGACA 2585

Qy       2748 TCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTT 2807
             |||||  ||  ||||||||  ||  ||  ||||||||||||||||  ||   ||||||||||||||
Db       2586 TCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCTGACAGGGAAATTT 2645

Qy       2808 ATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATT 2867
             |||||  ||||||||  ||  ||||||||||||||  ||||||||||||||||||  |||||||||
Db       2646 ATCACGCCAGAGCTTAACAATGTTGCCAGCCTCTGGTTCATGTCACTTTTCATCTGCATT 2705

Qy       2868 TTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGG 2927
             ||||||||||||||||  ||||||||||||||||||||  ||  ||  ||||||||||||||
Db       2706 TTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGATGACTGGTGGAGA 2765

Qy       2928 AATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGA 2987
             ||  ||||||||  ||||||||||||||  |||||  |||||  ||||||||||||||||||||
Db       2766 AACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGCTGTGTTCCAGGGA 2825

Qy       2988 CTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGAT 3047
             ||  ||||||||||||||||||||  ||  ||  ||||||||  ||||||||  |||||  |||||
Db       2826 CTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAGGGCGGAGAC 2885

Qy       3048 GATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACC 3107
             ||  ||||||||||||||||  ||  ||||||||||||||  |||  |  ||||||||||  ||
Db       2886 GACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCTGATACCTCCGACA 2945

Qy       3108 ACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAAC 3167
             |||  ||||  |||  ||||||||||||||  |||||  ||||||  |  ||  ||||||||||||  |||
Db       2946 ACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAAC 3005

Qy       3168 GGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTC 3227
             ||||||||  ||||||||||||||  ||  ||||||||  ||||||||||||||||||||  |||
Db       3006 GGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATTTTGGGTGATCGTC 3065

Qy       3228 CATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTC 3287
             ||||||||  ||  ||  |||||  |||  |||||||  |||||  |||||||  ||||||||||||
Db       3066 CATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTGTC 3125

Qy       3288 ATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCT 3347
             ||  ||||||||||  ||  ||||||||  ||||||||||  ||||||||  |||||  ||  ||
Db       3126 ATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGTCCGGATCGACCCG 3185

Qy       3348 TTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGAT 3407
             ||||||||||||||||||||||||  ||  |  ||||||||||||  ||||||||||||||||
Db       3186 TTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGATTGCAACTAGGAG 3245

Qy       3408 GTCAGTGCATCAGCTCCCCC 3427
             |||||  | |   || ||||
Db       3246 GTCAGCACGTGGACTTCCCC 3265
```

RESULT 6
AAZ58265
ID AAZ58265 standard; cDNA; 3786 BP.
XX
AC AAZ58265;
XX
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA clone p0097.cqrad17rc.
XX
KW Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key   Location/Qualifiers
FT CDS   2..3499
FT   /*tag= a
FT   /partial
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15871.
XX
PR 14-JUL-1998; 98US-0092844.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
DR WPI; 2000-182431/16.
DR P-PSDB; AAY58834.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
XX
PS Claim 9; Page 44-45; 93pp; English.
XX
CC The present sequence is that of cDNA clone p0097.cqrad17rc encoding
CC a portion (see AAY58833) of corn cellulose synthase (CS). The cDNA
CC clone was isolated from a european corn borer 4 times-infected corn
CC stage V9 whorl section cDNA library on the basis of homology to
CC Arabidopsis and cotton CS sequences. The invention relates to
CC isolated nucleic acid fragments encoding plant CS and to CS
CC polypeptides. It also relates to the construction of a chimeric
CC gene encoding all or a portion of the CS, in sense or antisense
CC orientation, where expression of the gene results in altered levels
CC of the CS in transformed host cells. The host cells can be used to
CC screen compounds for their ability to inhibit CS activity. CS
CC nucleic acids are also useful for producing transgenic plants having
CC altered levels of CS, and hence altered levels of fibre. CS may
CC also serve as a target for the development of novel herbicides.
XX
SQ Sequence 3786 BP; 846 A; 967 C; 1111 G; 862 T; 0 other;

  Query Match   56.9%; Score 2241; DB 21; Length 3786;
  Best Local Similarity 81.2%; Pred. No. 0;
  Matches 2691; Conservative   0; Mismatches 580; Indels   42; Gaps   6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
```


Db 194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253
Qy 183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
Db 254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313
Qy 242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
Db 314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373
Qy 297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
Db 374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433
Qy 357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
Db 434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493
Qy 417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
Db 494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553
Qy 477 GAGTTCAACTGG---------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
Db 554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613
Qy 522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
Db 614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664
Qy 582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
Db 665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724
Qy 642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
Db 725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784
Qy 693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
Db 785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844
Qy 753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
Db 845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904
Qy 813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
Db 905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964
Qy 870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
Db 965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024
Qy 930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
Db 1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084
Qy 990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
Db 1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144
Qy 1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109
Db 1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204
Qy 1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
Db 1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264
Qy 1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
Db 1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324
Qy 1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
Db 1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384
Qy 1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
Db 1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444
Qy 1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
Db 1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504
Qy 1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
Db 1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564
Qy 1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
Db 1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624
Qy 1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
Db 1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684
Qy 1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
Db 1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744
Qy 1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709
Db 1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804
Qy 1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769
Db 1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864
Qy 1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829
Db 1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924
Qy 1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889
Db 1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984
Qy 1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949
Db 1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044
Qy 1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009
Db 2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104
Qy 2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069
Db 2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164
Qy 2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129
Db 2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224
Qy 2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189
Db 2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284
Qy 2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249
Db 2285 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344
Qy 2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309
Db 2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

```
Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369
             ||| |||| |||||||||||| || ||||||||||||||  |||| |||||||| |||||
Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429
             || || ||| ||||||| |||||||| ||||||||| |||| ||||||||||||| ||
Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489
             || || || || |||||||||| |||||||||||||||||||||||| || |||||
Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549
              | ||||| ||||| | ||||| || |||||||| || ||||| |||||||| ||||||
Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609
             ||||| || || ||| |||||||||||||||||| |||||||| ||||| || |||||||
Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669
             |||||||| ||||| ||||| |||||||  | || || |||||||||| || |||||
Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729
             |||||||| || || || ||||| || |||||  ||||||| ||||| || |||||||||
Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
             ||||| ||||| ||||| || ||||| ||| |||||||||||| ||||||||||||||
Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
              |  | || ||||| |||||||| |||||||| |  ||||| ||||| | |||||||||
Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
             ||||||| |||||||| |  |  ||  ||||||| |||||||| |||||||| || |
Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
             || || |||||||||||||| |||||||||||||||| ||||| || ||  | || ||
Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
             || || |||||||||| || || |||||  | || ||  | || || ||||||||||||
Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
             || || ||||  || || || |||||||||| ||||| || || ||||| ||||| |||
Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
              | ||||||| || |||||  | || ||  |||||||||| || ||||| || ||  |
Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
             || || |||||||| ||||| || ||||| |||||||||| || |||||||| |||||
Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
             || || |||||||| ||||| || || || || ||||| ||| |||| || ||||| |||
Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
             ||||| || ||||| |||||||||||||||| |||||||| || ||||||||||||||
Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
             ||||| ||  | || || ||||| || |||    | || |||||| || ||||||||||
Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402
                ||||  |||||||
Db       3485 CTGGACTGCAACT 3497
```

RESULT 7
AAZ99509
ID AAZ99509 standard; DNA; 3813 BP.
XX
AC AAZ99509;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key   Location/Qualifiers
FT CDS   215..3493
FT   /*tag= a
FT   /product= "cellulose synthase"
FT   /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84113.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 1; Page 129-134; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 other;

Query Match 56.9%; Score 2241; DB 21; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;

Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy      123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
            ||  | || |||| |   | ||||||||| ||||| |||||||||||||||||| || |||
Db      194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy      183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
            || ||||||| || || |||||||||||| |||| ||  | ||||  | || ||  |
Db      254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy      242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
            ||    |||||       | |||||||| |||||||||| |||||  | | |  || |
Db      314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy      297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
            ||||| |||||||||||||||||||||||||||||||||||||| | |||||||||
Db      374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy      357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
            |||||||||| |||||| ||||    ||||| ||||||| ||| || ||||||||||||
Db      434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy      417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
            ||||||||||  || ||| |||| |||||||||||||||||||||| ||||||||   |
Db      494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy      477 GAGTTCAACTGG--------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
            ||||||  |  |                |||   | ||| ||||||  || |||||||||
Db      554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy      522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
            ||||||   |  || |||||||||||||| || || ||         |||||  |  |
Db      614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy      582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
            |||   | |||||| || || |||||||||||||| || |||||  ||||||||||||
Db      665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy      642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
            |||||||||||||||||| ||||| || | ||||         || || ||||| ||||||
Db      725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy      693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
            || || || || ||||  || |||||||  | |  || ||||||||  || || ||||||||
Db      785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy      753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
            |||||||||| || || || || ||||| || || || ||||||||   | ||||||
Db      845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy      813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
            ||||||||||  || ||| |  |||| ||     ||| || ||||| ||||||    |||
Db      905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy      870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
            || |||||||| ||||| |||||||||||||||| || || | | |||||| ||| |
Db      965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy      930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
            ||  |  |||||||||  |||||||| || |||||||||||| |||| ||| | ||||||
Db     1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084

Qy      990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
            ||||| |||||||||||||||||||||||||||||||| ||| ||||||||||| ||||||
Db     1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy     1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109
            |||||||||| |||||||||||||||||||||| ||||| |||||||||||||| ||||||
Db     1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy     1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
            |||||| | || || |||||||||||||||||| |||||  ||||||| ||||| ||||||
Db     1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy     1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
            ||||| || || ||||| |||||||||| || ||||||||||||||||||||||||| |
Db     1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy     1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
            |||||||||| ||||||||| | || ||||| || || |||||||| ||||||||||
Db     1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy     1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
            ||||| ||||| || ||||||||||||||||||||||||||||||| ||||||||||||
Db     1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy     1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
            || |||||||||||||||||||||||||||||||||||||| ||||| || |||||
Db     1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy     1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
            || ||||| ||||| |||||||||||||||||||||||||||||| ||||||||||||
Db     1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy     1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
            || ||    |||||||||||||||| || |||||||||| || ||||||||||||||
Db     1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy     1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
            || |||||||||| ||||| |||||||| || |||||||| ||||||||||||||||||
Db     1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy     1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
            ||||||| |||||||||||||||||| || || ||||||||||||||||||||||| |||
Db     1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy     1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709
            |||||||||||| || | |||    || |||||||| |||| ||  ||||||||||
Db     1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy     1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769
            || ||||| ||  | ||||| |||||||||| ||||| |||||||| ||||||||| ||
Db     1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy     1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829
            ||||| ||||||||||   || |||||||| ||||  | || |||||||| |||||||||
Db     1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy     1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889
            |||||||||| |||||||| ||||| || || |||||||| |||||||| |||||   |
Db     1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy     1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949
            || ||||| || |||||||||||||||||| || |||||||| |||||     ||||
Db     1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy     1950 CGATATGCTAACCGGAATGTTGTCTTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009
            ||||| |||||| |||| |||||||||||||| |||||||||||||||||||| ||||||
Db     2045 CGATACGCTAACAGGAACGTTGTCTTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy     2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069
            || || || |||||||| |||||||||||||| || ||| |||||||| | |||||||||
Db     2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy     2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129
            ||||| || ||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db     2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy     2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189
            | |  || ||||||    | || | |||  |||||||||||||| |||||| || |||
Db     2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy     2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAAACCAATCCCCTGCATATGCTCTTGGT 2249
            ||||| ||   |||||||||||||||| ||| |||| ||| |||||||||||| | |||
Db     2285 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344
```

```
Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309
             |||||||| |||| ||||||||| ||||| |  || |||||||| || ||||||||||
Db       2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369
             ||| |||| |||||||||||| || ||||||||||||||  |||| |||||||| |||||
Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429
             || || ||| ||||||| |||||||| ||||||||| |||| |||||||||||||| ||
Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489
             || || || || |||||||||| |||||||||||||||||||||||||| || |||||
Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549
              | ||||| |||||  | ||||| || |||||||| || ||||| |||||||| ||||||
Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609
             ||||| || || |||  |||||||||||||||||| |||||||| ||||| || ||||||
Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669
             |||||||| ||||| ||||| |||||||| | || || |||||||||| || |||||
Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729
             |||||||| || || || ||||| || |||||  |||||||| ||||| || |||||||||
Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
             ||||| ||||| ||||| || ||||| ||| |||||||||||| ||||||||||||||
Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
              |  | || ||||| |||||||| |||||||  | |||||| |||||  | |||||||||
Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
              ||||||| |||||||| |  |  ||  ||||||| |||||||| |||||||| || |
Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
             || || |||||||||||||| |||||||||||||||| ||||| || ||  | || ||
Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
             || || |||||||||| || || |||||  | || ||  | || || ||||||||||||
Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
             || || ||||  || || || |||||||||| ||||| || || ||||| ||||| |||
Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
              |  ||||||| || |||||  | || ||  |||||||||| || ||||| || ||  |
Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
             || || |||||||| ||||| || ||||| |||||||||| || |||||||| |||||
Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
             || || |||||||| ||||| || || || || ||||| ||| |||| || ||||| |||
Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
             ||||| || ||||| |||||||||||||||| |||||||| || ||||||||||||||
Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
```

```
                  ||||| ||  | || || ||||| || |||    | || ||||| || ||||||||||||
Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402
                  |||| |||||||
Db       3485 CTGGACTGCAACT 3497
```

```
RESULT 8
AAZ99524
ID   AAZ99524 standard; DNA; 3813 BP.
XX
AC   AAZ99524;
XX
DT   03-JUL-2000  (first entry)
XX
DE   DNA encoding a maize cellulose synthase.
XX
KW   Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW   transgenic plant; plant breeding marker; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             215..3493
FT                   /*tag=  a
FT                   /product= "cellulose synthase"
FT                   /note= "no termination codon given"
XX
PN   WO200009706-A2.
XX
PD   24-FEB-2000.
XX
PF   16-AUG-1999;   99WO-US18760.
XX
PR   17-AUG-1998;   98US-0096822.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Dhugga KS,  Helentjaris TG,  Bowen BA,  Wang X;
XX
DR   WPI; 2000-224343/19.
DR   P-PSDB; AAY84118.
XX
PT   New genes which encode maize cellulose synthase polypeptides in plants
PT   useful for modulating the expression of cellulose synthase in plants
PT   and to produce transgenic plants expressing the novel protein  -
XX
PS   Claim 1; Page 168-173; 119pp; English.
XX
CC   The present sequence encodes a maize cellulose synthase polypeptide.
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 other;
```

Query Match 56.9%; Score 2241; DB 21; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| ||||||||||||||||||| || |||
Db        194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || |||||||| || || |||||||||||| |||| ||  | ||||  | || ||  |
Db        254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||    |||||       | ||||||||| |||||||||| |||||  | | |  || |
Db        314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||| | |||||||||
Db        374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||| |||||| ||||    |||||| ||||||| ||| || ||||||||||||
Db        434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              ||||||||||  || ||| |||| |||||||||||||||||||||| ||||||||   |
Db        494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy        477 GAGTTCAACTGG---------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              ||||||  | |                |||  | ||| |||||  || |||||||||
Db        554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
              ||||||   |  || |||||||||||||| || || ||          |||||  | |
Db        614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
              |||   | |  |||||| || || |||||||||||| || |||||| ||||||||||
Db        665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
              |||||||||||||||| ||||| || | ||||          || || ||||| ||||||
Db        725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
              || || || || ||||  || |||||||  | | || |||||||| || || ||||||||
Db        785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
              |||||||||| || || || || ||||| || || || ||||||||  | ||||||
Db        845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
              ||||||||||  || ||| |  |||| ||     |||  || ||||| ||||||  . |||
Db        905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
              || |||||||| ||||| |||||||||||||||| || || | | ||||||| ||| |
Db        965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
              ||  |  |||||||||  |||||||| || |||||||||||| |||| ||| | ||||||
Db       1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084

Qy        990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
              ||||| |||||||||||||||||||||||||||||||| ||| ||||||||||| ||||||
Db       1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy       1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109
              |||||||||| |||||||||||||||||||| ||||| |||||||||||||| ||||||
Db       1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy       1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
              ||||||  | ||  || |||||||||||||||||| |||||  |||||||  |||||  ||||||
Db       1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy       1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
              |||||  ||  ||  |||||  |||||||||||  ||  |||||||||||||||||||||||||  |
Db       1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy       1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
              ||||||||||  |||||||||  | ||  |||||  ||  ||  ||||||||  ||||||||||
Db       1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy       1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
              |||||  |||||  ||  ||||||||||||||||||||||||||||||  ||||||||||
Db       1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy       1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
              ||  ||||||||||||||||||||||||||||||||||||||||  |||||  ||  |||||
Db       1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy       1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
              ||  |||||  |||||  ||||||||||||||||||||||||||||  ||||||||||||
Db       1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy       1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
              ||  ||     ||||||||||||||||||  |||||||||||  ||  ||||||||||||||
Db       1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy       1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
              ||  ||||||||||  |||||  ||||||||  ||  ||||||||  |||||||||||||||
Db       1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy       1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
              |||||||  ||||||||||||||||||  ||  ||  ||||||||||||||||||||||  |||
Db       1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy       1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709
              ||||||||||| ||  ||  | |||     ||  ||||||||  ||||  ||  ||||||||||
Db       1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy       1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769
              ||  |||||  ||   | |||||  ||||||||||  |||||  ||||||||  |||||||||  ||
Db       1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy       1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829
              |||||  ||||||||||   ||  ||||||||  ||||   | ||  ||||||||  |||||||||
Db       1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy       1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889
              ||||||||||  ||||||||  |||||  ||  ||  ||||||||  ||||||||  |
Db       1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy       1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949
              ||  |||||  ||  ||||||||||||||||||||  ||  ||||||||  |||||      ||||
Db       1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy       1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009
              |||||  ||||||  ||||  ||||||||||||||  ||||||||||||||||||||  ||||||
Db       2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy       2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069
              ||  ||  ||  ||||||||  ||||||||||||||  ||  |||  ||||||||  | |||||||||
Db       2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy       2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129
              |||||  ||  |||||  ||||||||||||||||||||  ||||||||||||||||||||||||
Db       2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy       2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189
              | |  || ||||||    | || | |||  ||||||||||||| |||||  || || |||
Db       2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy       2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249
```

```
Db       2285 AAGAAGAAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344

Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309

Db       2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369

Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429

Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489

Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549

Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609

Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669

Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729

Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789

Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849

Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909

Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969

Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029

Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089

Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149

Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209

Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269

Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329

Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389

Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402

Db       3485 CTGGACTGCAACT 3497
```

RESULT 9
ABZ14725
ID ABZ14725 standard; DNA; 3255 BP.
XX
AC ABZ14725;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2530.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-264647P.
PR 22-JUN-2001; 2001US-300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT -
XX
PS Claim 144; SEQ ID NO 2530; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 3255 BP; 841 A; 641 C; 813 G; 960 T; 0 other;

Query Match 38.3%; Score 1509; DB 24; Length 3255;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 2207; Conservative 0; Mismatches 1030; Indels 30; Gaps 5;

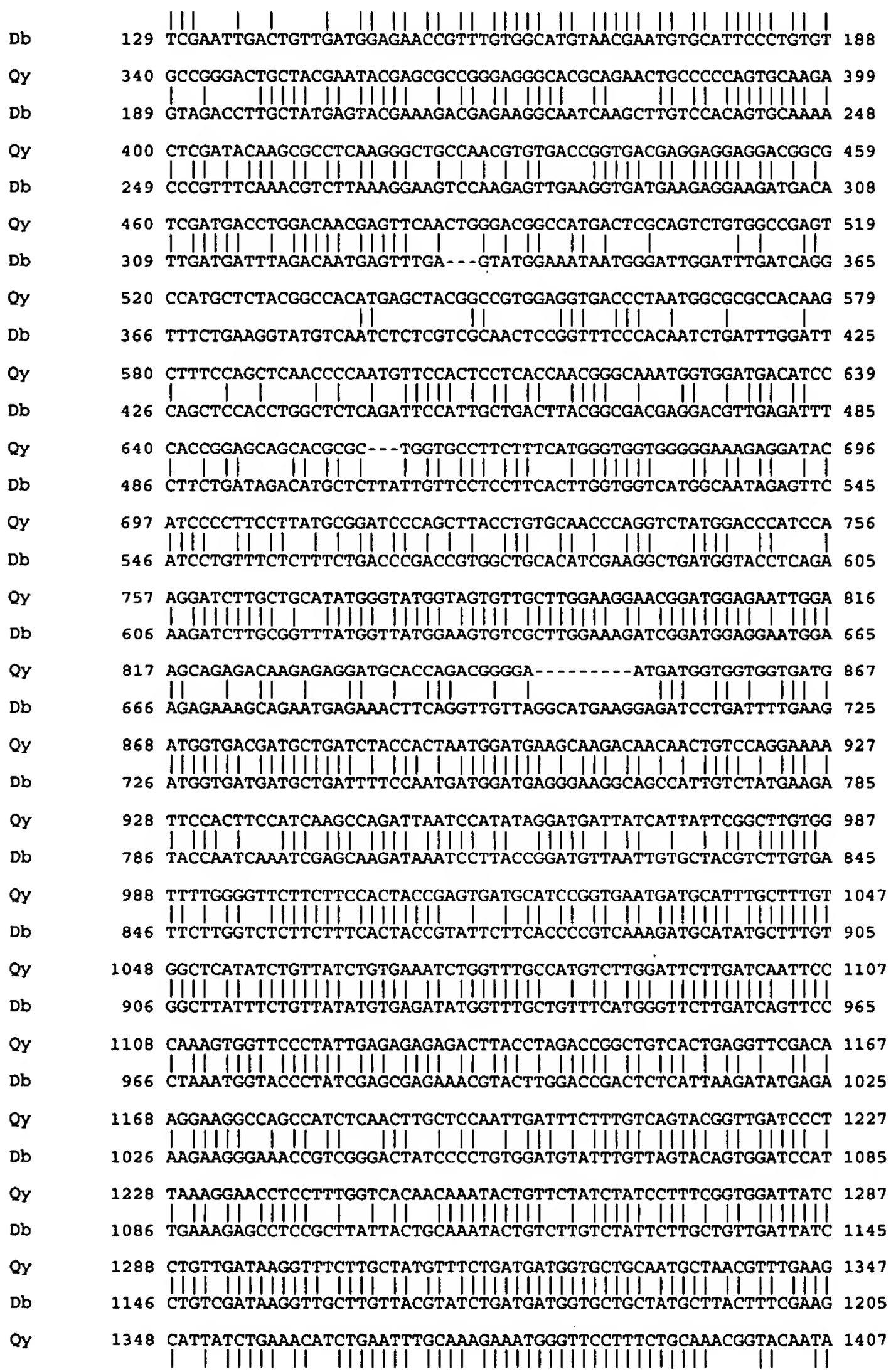
```
Qy        144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203

Db          1 ATGAACACCGGTGGTCGGTTAATCGCCGGTTCTCACAACAGGAATGAGTTTGTCCTCATT 60
```

```
Qy       204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db        61 AATGCCGATGAGAATGCCCGAATAAGATCAGTCCAAGAGCTGAGTGGACAGACATGTCAA 120

Qy       264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db       121 ATCTGCAGAGATGAGATCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAACGAA 180

Qy       324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db       181 TGTGCATTCCCTGTGTGTAGACCTTGCTATGAGTACGAAAGACGAGAAGGCAATCAAGCT 240

Qy       384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
Db       241 TGTCCACAGTGCAAAACCCGTTTCAAACGTCTTAAAGGAAGTCCAAGAGTTGAAGGTGAT 300

Qy       444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
Db       301 GAAGAGGAAGATGACATTGATGATTTAGACAATGAGTTTGA---GTATGGAAATAATGGG 357

Qy       504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db       358 ATTGGATTTGATCAGGTTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCCA 417

Qy       564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db       418 CAATCTGATTTGGATTCAGCTCCACCTGGCTCTCAGATTCCATTGCTGACTTACGGCGAC 477

Qy       624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGC---TGGTGCCTTCTTTCATGGGTGGT 680
Db       478 GAGGACGTTGAGATTTCTTCTGATAGACATGCTCTTATTGTTCCTCCTTCACTTGGTGGT 537

Qy       681 GGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGG 740
Db       538 CATGGCAATAGAGTTCATCCTGTTTCTCTTTCTGACCCGACCGTGGCTGCACATCCAAGG 597

Qy       741 TCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAA 800
Db       598 CCTATGGTACCTCAGAAAGATCTTGCGGTTTATGGTTATGGAAGTGTCGCTTGGAAAGAT 657

Qy       801 CGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGA---------AT 851
Db       658 CGGATGGAGGAATGGAAGAGAAAGCAGAATGAGAAACTTCAGGTTGTTAGGCATGAAGGA 717

Qy       852 GATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAA 911
Db       718 GATCCTGATTTTGAAGATGGTGATGATGCTGATTTTCCAATGATGGATGAGGGAAGGCAG 777

Qy       912 CAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATC 971
Db       778 CCATTGTCTAGGAAGATACCAATCAAATCGAGCAAGATAAATCCTTACCGGATGTTAATT 837

Qy       972 ATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAAT 1031
Db       838 GTGCTACGTCTTGTGATTCTTGGTCTCTTCTTTCACTACCGTATTCTTCACCCCGTCAAA 897

Qy      1032 GATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGG 1091
Db       898 GATGCATATGCTTTGTGGCTTATTTCTGTTATATGTGAGATATGGTTTGCTGTTTCATGG 957

Qy      1092 ATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTG 1151
Db       958 GTTCTTGATCAGTTCCCTAAATGGTACCCTATCGAGCGAGAAACGTACTTGGACCGACTC 1017

Qy      1152 TCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTC 1211
Db       1018 TCATTAAGATATGAGAAAGAAGGGAAACCGTCGGGACTATCCCCTGTGGATGTATTTGTT 1077

Qy      1212 AGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATC 1271
Db       1078 AGTACAGTGGATCCATTGAAAGAGCCTCCGCTTATTACTGCAAATACTGTCTTGTCTATT 1137

Qy      1272 CTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCA 1331
```

```
Db      1138 CTTGCTGTTGATTATCCTGTCGATAAGGTTGCTTGTTACGTATCTGATGATGGTGCTGCT 1197

Qy      1332 ATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTC 1391
Db      1198 ATGCTTACTTTCGAAGCTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTC 1257

Qy      1392 TGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGAC 1451
Db      1258 TGCAAGAAATATTGTATTGAGCCTCGTGCTCCCGAATGGTATTTCTGCCATAAAATGGAC 1317

Qy      1452 TACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAG 1511
Db      1318 TACTTGAAGAATAAAGTTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGAT 1377

Qy      1512 TATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAA 1571
Db      1378 TATGAAGAATTCAAAGTAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCCTGAG 1437

Qy      1572 GAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCT 1631
Db      1438 GATGGTTGGACTATGCAAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCCT 1497

Qy      1632 GGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTG 1691
Db      1498 GGCATGATTCAGGTCTTCCTTGGAAGTGACGGTGTTCGTGATGTCGAAAACAACGAGTTG 1557

Qy      1692 CCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCT 1751
Db      1558 CCTCGATTAGTTTACGTTTCTCGTGAGAAGAGACCCGGATTTGATCACCATAAGAAGGCT 1617

Qy      1752 GGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTA 1811
Db      1618 GGAGCTATGAATTCCCTGATACGAGTCTCTGGGGTTCTATCAAATGCTCCTTACCTTCTG 1677

Qy      1812 AACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTT 1871
Db      1678 AATGTCGATTGTGATCACTACATCAACAATAGCAAAGCTCTTAGAGAAGCAATGTGTTTC 1737

Qy      1872 ATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGAT 1931
Db      1738 ATGATGGATCCTCAGTCAGGAAAGAAAATCTGTTATGTTCAGTTCCCTCAAAGGTTCGAT 1797

Qy      1932 GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATG 1991
Db      1798 GGGATTGATAGGCACGATCGATACTCAAATCGCAATGTTGTGTTCTTTGATATCAATATG 1857

Qy      1992 AAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGG 2051
Db      1858 AAAGGTTTGGATGGGCTACAAGGGCCTATATACGTCGGTACAGGTTGTGTTTTCAGGAGG 1917

Qy      2052 CAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAAC 2111
Db      1918 CAAGCGCTTTACGGATTTGATGCACCGAAGAAGAAGAAGGGCCCACGTAAGACATGCAAT 1977

Qy      2112 TGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACT 2171
Db      1978 TGCTGGCCAAAATGGTGTCTCCTATGTTTTGGTTCAAGAAAGAATCGTAAAGCAAAGACA 2037

Qy      2172 ACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCC 2231
Db      2038 GTGGCTGCGGATAAGAAGAAGAAGAA----------TAGGGAAGCGTCAAAGCAGATC 2085

Qy      2232 CCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCC 2291
Db      2086 CACGCATTAGAAAATATCGAAGAGGGCCGCGTCACTAAAGGTTCTAACGTAGAACAGTCA 2145

Qy      2292 GGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACA 2351
Db      2146 ACCGAGGCAATGCAAATGAAGTTGGAGAAGAAATTTGGGCAGTCTCCTGTATTTGTTGCA 2205

Qy      2352 TCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAA 2411
```

```
Db       2206 TCTGCGCGTATGGAGAATGGTGGGATGGCTAGAAACGCAAGCCCGGCTTGTCTGCTTAAA 2265

Qy       2412 GAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATT 2471
             ||||| || || |||||||||||| || |||||||| || || || ||||||||||||||||
Db       2266 GAAGCCATCCAAGTCATTAGTTGCGGATATGAAGATAAAACTGAATGGGGAAAAGAGATT 2325

Qy       2472 GGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCAT 2531
             || |||||||||| || ||||| |||||||||| || ||||| ||||||||| ||||
Db       2326 GGGTGGATCTATGGTTCTGTTACCGAAGATATTCTTACGGGTTTTAAGATGCATTCTCAT 2385

Qy       2532 GGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTG 2591
             |||||| | ||  |||| || | ||| ||    |  || |||||||| || || ||  |
Db       2386 GGTTGGAGATCTGTTTATTGTACACCAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATC 2445

Qy       2592 AATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651
             |||||||| |||||||| || || || ||||| ||||| ||||||||  ||||||| |||
Db       2446 AATCTTTCGGATCGTCTCCATCAAGTTCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTC 2505

Qy       2652 TTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGA 2711
             || || |  ||||| ||| |||||||||| ||||| || ||  ||||||   | ||  |
Db       2506 TTGAGTAGGCATTGTCCTATTTGGTATGGTTATGGAGGTGGGTTGAAATGGCTTGAGCGG 2565

Qy       2712 TTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGT 2771
             || ||||||| |||||  | || || || ||||| ||| | || ||| | | |||||||
Db       2566 TTGTCCTACATTAACTCTGTGGTTTACCCGTGGACCTCTCTACCGCTCATCGTTTACTGT 2625

Qy       2772 ACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTT 2831
               | | |||||||||||| |  | || || ||||| |||  ||| ||| | |  ||   |
Db       2626 TCTCTCCCTGCCATCTGTCTTCTCACTGGAAAATTCATCGTTCCCGAGATTAGCAACTAT 2685

Qy       2832 GCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATG 2891
             || ||  |    |||||| | || ||    |  |||      ||| | || || || |||
Db       2686 GCGAGTATCCTCTTCATGGCGCTCTTCTCGTCGATTGCAATAACGGGTATTCTCGAGATG 2745

Qy       2892 AGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGA 2951
               ||||  |    ||||| || ||||| |||||||| || || ||||| ||||||||||||
Db       2746 CAATGGGGCAAAGTTGGGATCGATGATTGGTGGAGAAACGAACAGTTTTGGGTCATTGGA 2805

Qy       2952 GGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTT 3011
             ||||| ||  | || || |||||| | ||||| || || ||||||||  | ||||||||
Db       2806 GGTGTTTCTGCGCATCTGTTTGCTCTCTTCCAAGGTCTCCTCAAGGTTCTTGCTGGTGTC 2865

Qy       3012 GATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
             || || | |||||| || |||||||| |  |  |||||||  |||||||| || || ||
Db       2866 GACACTAACTTCACAGTCACATCAAAAGCAGCTGATGATGGAGAGTTCTCTGACCTTTAC 2925

Qy       3072 ACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATT 3131
               |||||||||||| |  | || |||||  |  ||  |  ||  |  | ||| |||||
Db       2926 CTCTTCAAATGGACTTCACTTCTCATCCCTCCAATGACTCTACTCATCATAAACGTCATT 2985

Qy       3132 GGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTC 3191
             || ||  | |  || || ||  |||| ||||  || ||||| || || ||||| || ||
Db       2986 GGAGTCATAGTCGGAGTCTCTGATGCCATCAGCAATGGATACGACTCGTGGGGACCGCTT 3045

Qy       3192 TTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGT 3251
             || || |  || ||||||||| ||||||| ||  | |||||||| || || || ||||||
Db       3046 TTCGGAAGACTGTTCTTTGCACTTTGGGTCATCATTCATCTTTACCCGTTCCTTAAAGGT 3105

Qy       3252 TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCT 3311
             ||| |||| |  ||| | || |  ||||| ||| |  |||||||||||| || |||||
Db       3106 TTGCTTGGGAAACAAGATAGAATGCCAACCATTATTGTCGTCTGGTCCATCCTCCTGGCC 3165

Qy       3312 TCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCG 3371
             || ||  |  | || |||||||| ||| || |||| ||  | ||||   | |  |||||
Db       3166 TCGATTCTTACACTTCTTTGGGTCCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCCT 3222

Qy       3372 CTTCTTGAGGAGTGTGGTTTGGATTGC 3398
               |||| |||   |||||||| || |||
Db       3223 ATTCTCGAGATCTGTGGTTTAGACTGC 3249


RESULT 10
AAC83798
```

```
ID   AAC83798 standard; cDNA; 3444 BP.
XX
AC   AAC83798;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Arabidopsis thaliana cellulose synthase CelA cDNA.
XX
KW   Arabidopsis thaliana; aspen; CelA; cellulose synthase; wood;
KW   mechanical stress response element; MSRE; cellulose biosynthesis; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200071670-A2.
XX
PD   30-NOV-2000.
XX
PF   18-MAY-2000; 2000WO-US13637.
XX
PR   21-MAY-1999;   99US-0135280.
XX
PA   (UNMT ) UNIV MICHIGAN TECHNOLOGICAL.
XX
PI   Chiang VL,  Wu L,  Chandrashekhar JP,  Carraway DT;
XX
DR   WPI; 2001-032025/04.
XX
PT   Novel polynucleotide encoding cellulose synthase useful for enhancing
PT   cellulose biosynthesis, modifying lignin biosynthesis, and for
PT   accelerating growth of plants and for upregulating cellulose synthase
PT   levels -
XX
PS   Claim 8; Fig 7; 65pp; English.
XX
CC   The present sequence is given in a specification relating to
CC   polynucleotides encoding cellulose synthase enzymes from Populus
CC   tremuloides and Arabidopsis. The cellulose synthases are designated
CC   PtCelA and AraxCelA respectively. The polynucleotides may be introduced
CC   into plants where they are useful for altering the growth of a plant.
CC   Vectors comprising a cellulose synthase promoter, or a fragment of the
CC   promoter containing one or more mechanical stress response elements
CC   (MSRE), operably linked to a cellulose synthase gene are useful for
CC   causing stress-induced gene expression in a plant cell. A cDNA encoding a
CC   protein that binds to a positive MSRE may be delivered into the plant to
CC   increase expression of cellulose synthase. A cDNA in an antisense
CC   orientation may be delivered to reduce expression of cellulose synthase.
CC   The polynucleotides enable the development of transgenic tree varieties
CC   having increased cellulose content, decreased lignin content and
CC   therefore improved wood fibre characteristics. The cellulose synthase
CC   promoters are useful for synthesising greater quantities of
CC   high-crystalline cellulose in plant and preferably in trees, which
CC   permits an increase in proportion of cellulose in transgenic plants,
CC   greater strength of juvenile wood and fiber and acceleration of overall
CC   growth rate.
CC   Note: This nucleotide sequence is given in the specification as
CC   SEQ ID NO: 5. SEQ ID NO: 5 is referred to in claim 8 but it is described
CC   as a polypeptide sequence.
XX
SQ   Sequence 3444 BP; 903 A; 664 C; 860 G; 1017 T; 0 other;

  Query Match             37.9%;  Score 1489.8;  DB 22;  Length 3444;
  Best Local Similarity   67.3%;  Pred. No. 0;
  Matches 2189;  Conservative    0;  Mismatches 1032;  Indels   30;  Gaps    5;

Qy        160 GGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATC 219
              || |  | ||||| || |||||| | || ||| | ||| ||||      ||| |   ||
Db          9 GGTTAATCGCCGGTTCTCACAACAGGAATGAGTTTGTCCTCATTAATGCCGATGAGAATG 68

Qy        220 CCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACG 279
              || |    |   |  | |||| ||  || |||   || || || ||| | || ||
Db         69 CCCGAATAAGATCAGTCCAAGAGCTGAGTGGACAGACATGTCAAATCTGCAGAGATGAGA 128

Qy        280 TCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCT 339
```


Db 129 TCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAACGAATGTGCATTCCCTGTGT 188
Qy 340 GCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGA 399
Db 189 GTAGACCTTGCTATGAGTACGAAAGACGAGAAGGCAATCAAGCTTGTCCACAGTGCAAAA 248
Qy 400 CTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCG 459
Db 249 CCCGTTTCAAACGTCTTAAAGGAAGTCCAAGAGTTGAAGGTGATGAAGAGGAAGATGACA 308
Qy 460 TCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGT 519
Db 309 TTGATGATTTAGACAATGAGTTTGA---GTATGGAAATAATGGGATTGGATTTGATCAGG 365
Qy 520 CCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAG 579
Db 366 TTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCCACAATCTGATTTGGATT 425
Qy 580 CTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCC 639
Db 426 CAGCTCCACCTGGCTCTCAGATTCCATTGCTGACTTACGGCGACGAGGACGTTGAGATTT 485
Qy 640 CACCGGAGCAGCACGCGC---TGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATAC 696
Db 486 CTTCTGATAGACATGCTCTTATTGTTCCTCCTTCACTTGGTGGTCATGGCAATAGAGTTC 545
Qy 697 ATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCA 756
Db 546 ATCCTGTTTCTCTTTCTGACCCGACCGTGGCTGCACATCGAAGGCTGATGGTACCTCAGA 605
Qy 757 AGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGA 816
Db 606 AAGATCTTGCGGTTTATGGTTATGGAAGTGTCGCTTGGAAAGATCGGATGGAGGAATGGA 665
Qy 817 AGCAGAGACAAGAGAGGATGCACCAGACGGGGA---------ATGATGGTGGTGGTGATG 867
Db 666 AGAGAAAGCAGAATGAGAAACTTCAGGTTGTTAGGCATGAAGGAGATCCTGATTTTGAAG 725
Qy 868 ATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAA 927
Db 726 ATGGTGATGATGCTGATTTTCCAATGATGGATGAGGGAAGGCAGCCATTGTCTATGAAGA 785
Qy 928 TTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGG 987
Db 786 TACCAATCAAATCGAGCAAGATAAATCCTTACCGGATGTTAATTGTGCTACGTCTTGTGA 845
Qy 988 TTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGT 1047
Db 846 TTCTTGGTCTCTTCTTTCACTACCGTATTCTTCACCCCGTCAAAGATGCATATGCTTTGT 905
Qy 1048 GGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCC 1107
Db 906 GGCTTATTTCTGTTATATGTGAGATATGGTTTGCTGTTTCATGGGTTCTTGATCAGTTCC 965
Qy 1108 CAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACA 1167
Db 966 CTAAATGGTACCCTATCGAGCGAGAAACGTACTTGGACCGACTCTCATTAAGATATGAGA 1025
Qy 1168 AGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCT 1227
Db 1026 AAGAAGGGAAACCGTCGGGACTATCCCCTGTGGATGTATTTGTTAGTACAGTGGATCCAT 1085
Qy 1228 TAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATC 1287
Db 1086 TGAAAGAGCCTCCGCTTATTACTGCAAATACTGTCTTGTCTATTCTTGCTGTTGATTATC 1145
Qy 1288 CTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAG 1347
Db 1146 CTGTCGATAAGGTTGCTTGTTACGTATCTGATGATGGTGCTGCTATGCTTACTTTCGAAG 1205
Qy 1348 CATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATA 1407
Db 1206 CTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTCTGCAAGAAATATTGTA 1265
Qy 1408 TTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGG 1467
Db 1266 TTGAGCCTCGTGCTCCCGAATGGTATTTCTGCCATAAAATGGACTACTTGAAGAATAAAG 1325
Qy 1468 TGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGG 1527
Db 1326 TTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGATTATGAAGAATTCAAAG 1385
Qy 1528 TGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGC 1587
Db 1386 TAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCCTGAGGATGGTTGGACTATGC 1445
Qy 1588 AAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCT 1647
Db 1446 AAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCCTGGCATGATTCAGGTCT 1505
Qy 1648 TCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATG 1707
Db 1506 TCCTTGGAAGTGACGGTGTTCGTGATGTCGAAAACAACGAGTTGCCTCGATTAGTTTACG 1565
Qy 1708 TTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCAT 1767
Db 1566 TTTCTCGTGAGAAGAGACCCGGATTTGATCACCATAAGAAGGCTGGAGCTATGAATTCCC 1625
Qy 1768 TGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATC 1827
Db 1626 TGATACGAGTCTCTGGGGTTCTATCAAATGCTCCTTACCTTCTGAATGTCGATTGTGATC 1685
Qy 1828 ACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTAC 1887
Db 1686 ACTACATCAACAATAGCAAAGCTCTTAGAGAAGCAATGTGTTTCATGATGGATCCTCAGT 1745
Qy 1888 TAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATG 1947
Db 1746 CAGGAAAGAAAATCTGTTATGTTCAGTTCCCTCAAAGGTTCGATGGGATTGATAGGCACG 1805
Qy 1948 ACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTA 2007
Db 1806 ATCGATACTCAAATCGCAATGTTGTGTTCTTTGATATCAATATGAAAGGTTTGGATGGGC 1865
Qy 2008 TTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTT 2067
Db 1866 TACAAGGGCCTATATACGTCGGTACAGGTTGTGTTTTCAGGAGGCAAGCGCTTTACGGAT 1925
Qy 2068 ATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGT 2127
Db 1926 TTGATGCACCGAAGAAGAAGAAGGGCCCACGTAAGACATGCAATTGCTGGCCAAAATGGT 1985
Qy 2128 GCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAG 2187
Db 1986 GTCTCCTATGTTTTGGTTCAAGAAAGAATCGTAAAGCAAAGACAGTGGCTGCGGATAAGA 2045
Qy 2188 AGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTG 2247
Db 2046 AGAAGAAGAA----------TAGGGAAGCGTCAAAGCAGATCCACGCATTAGAAAATA 2093
Qy 2248 GTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAAC 2307
Db 2094 TCGAAGAGGGCCGCGGTCATAAAGTTCTTAACGTAGAACAGTCAACCGAGGCAATGCAAA 2153
Qy 2308 AAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGA 2367
Db 2154 TGAAGTTGCAGAAGAAATATGGGCAGTCTCCTGTATTTGTTGCATCTGCGCGTCTGGAGA 2213
Qy 2368 ATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCA 2427
Db 2214 ATGGTGGGATGGCTAGAAACGCAAGCCCGGCTTGTCTGCTTAAAGAAGCCATCCAAGTCA 2273
Qy 2428 TTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGAT 2487
Db 2274 TTAGTCGCGGATATGAAGATAAAACTGAATGGGGAAAAGAGATTGGGTGGATCTATGGTT 2333

```
Qy       2488 CAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTT 2547
Db       2334 CTGTTACCGAAGATATTCTTACGGGTTCTAAGATGCATTCTCATGGTTGGAGACATGTTT 2393

Qy       2548 ACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTC 2607
Db       2394 ATTGTACACCAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATCAATCTTTCGGATCGTC 2453

Qy       2608 TTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCC 2667
Db       2454 TCCATCAAGTTCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTCTTGAGTAGGCATTGTC 2513

Qy       2668 CTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACT 2727
Db       2514 CTATTTGGTATGGTTATGGAGGTGGGTTGAAATGGCTTGAGCGGTTGTCCTACATTAACT 2573

Qy       2728 CCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCT 2787
Db       2574 CTGTGGTTTACCCGTGGACCTCTCTACCGCTCATCGTTTACTGTTCTCTCCCTGCCATCT 2633

Qy       2788 GTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCA 2847
Db       2634 GTCTTCTCACTGGAAAATTCATCGTTCCCGAGATTAGCAACTATGCGAGTATCCTCTTCA 2693

Qy       2848 TGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTG 2907
Db       2694 TGGCGCTCTTCTCGTCGATTGCAATAACGGGTATTCTCGAGATGCAATGGGGCAAAGTTG 2753

Qy       2908 GAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACC 2967
Db       2754 GGATCGATGATTGGTGGAGAAACGAACAGTTTTGGGTCATTGGAGGTGTTTCTGCGCATC 2813

Qy       2968 TCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCG 3027
Db       2814 TGTTTGCTCTCTTCCAAGGTCTCCTCAAGGTTCTTGCTGGTGTCGACACTAACTTCACAG 2873

Qy       3028 TGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTA 3087
Db       2874 TCACATCAAAAGCAGCTGATGATGGAGAGTTCTCTGACCTTTACCTCTTCAAATGGACTT 2933

Qy       3088 CCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCG 3147
Db       2934 CACTTCTCATCCCTCCAATGACTCTACTCATCATAAACGTCATTGGAGTCATAGTCGGAG 2993

Qy       3148 TTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCT 3207
Db       2994 TCTCTGATGCCATCAGCAATGGATACGACTCGTGGGGACCGCTTTTCGGAAGACTGTTCT 3053

Qy       3208 TTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAA 3267
Db       3054 TTGCACTTTGGGTCATCATTCATCTTTACCCGTTCCTTAAAGGTTTGCTTGGGAAACAAG 3113

Qy       3268 ACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCC 3327
Db       3114 ATAGAATGCCAACCATTATTGTCGTCTGGTCCATCCTCCTGGCCTCGATTCTTACACTTC 3173

Qy       3328 TTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTG 3387
Db       3174 TTTGGGTCCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCCTATTCTCGAGATCTGTG 3230

Qy       3388 GTTTGGATTGC 3398
Db       3231 GTTTAGACTGC 3241
```

RESULT 11
ABZ12754
ID ABZ12754 standard; DNA; 3255 BP.
XX
AC ABZ12754;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 559.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-264647P.
PR 22-JUN-2001; 2001US-300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT -
XX
PS Claim 144; SEQ ID NO 559; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 3255 BP; 884 A; 621 C; 798 G; 952 T; 0 other;

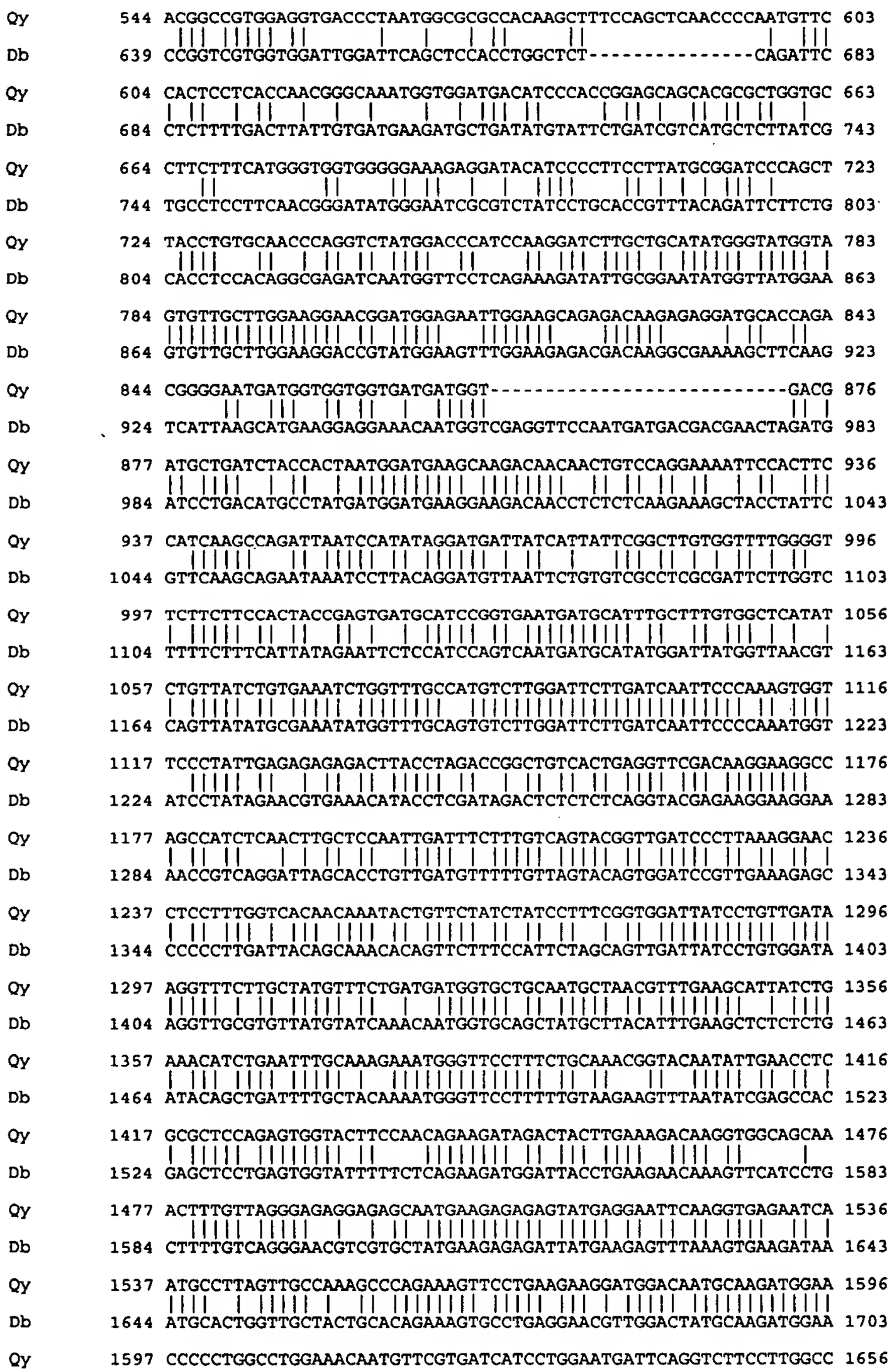
Query Match 36.4%; Score 1431.8; DB 24; Length 3255;
Best Local Similarity 66.0%; Pred. No. 0;
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

```
Qy        144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203
Db          1 ATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACAGAAACGAATTCGTTCTCATT 60

Qy        204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db         61 AACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAACTGAGTGGGCAAACATGTCAA 120

Qy        264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db        121 ATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGCTCTTTGTTGCTTGCAACGAA 180

Qy        324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db        181 TGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAACGTAGAGAAGGAAATCAAGCT 240

Qy        384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
Db        241 TGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTAGTCCACGGGTTGATGGAGAT 300

Qy        444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
```

```
Db        301 GATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGGATGGACCCT 360
Qy        504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db        361 GAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACACCGGTCGTGGTGGATTGGAT 420
Qy        564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db        421 TCAGCTCCACCTG----------------GCTCTCAGATTCCTCTTTTGACTTATTGTGAT 465
Qy        624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGG 683
Db        466 GAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCGTGCCTCCTTCAACGGGATAT 525
Qy        684 GGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCT 743
Db        526 GGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTGCACCTCCACAGGCGAGATCA 585
Qy        744 ATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGG 803
Db        586 ATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAAGTGTTGCTTGGAAGGACCGT 645
Qy        804 ATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGT 863
Db        646 ATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAGTCATTAAGCATGAAGGAGGA 705
Qy        864 GATGATGGT---------------------------GACGATGCTGATCTACCACTAATG 896
Db        706 AACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATGATCCTGACATGCCTATGATG 765
Qy        897 GATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCA 956
Db        766 GATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTCGTTCAAGCAGAATAAATCCT 825
Qy        957 TATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTG 1016
Db        826 TACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTCTTTTCTTTCATTATAGAATT 885
Qy       1017 ATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGG 1076
Db        886 CTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGTCAGTTATATGCGAGATATGG 945
Qy       1077 TTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACT 1136
Db        946 TTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGTATCCTATAGAACGTGAAACA 1005
Qy       1137 TACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCA 1196
Db       1006 TACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAAAACCGTCAGGATTAGCACCT 1065
Qy       1197 ATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAAT 1256
Db       1066 GTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGCCACCCTTGATTACAGCAAAC 1125
Qy       1257 ACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCT 1316
Db       1126 ACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATAAGGTTGCGTGTTATGTATCA 1185
Qy       1317 GATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAG 1376
Db       1186 GACGATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTGATACAGCTGAGTTTGCTAGA 1245
Qy       1377 AAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTC 1436
Db       1246 AAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCACGAGCTCCTGAGTGGTATTTT 1305
Qy       1437 CAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGA 1496
Db       1306 TCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTGCTTTTGTCAGGGAACGTCGT 1365
Qy       1497 GCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCC 1556
Db       1366 GCTATGAAGAGAGATTATGAGGAGTTTAAAGTGAAGATAAATGCACTGGTTGCTACTGCA 1425
```

```
Qy       1557 CAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAAT 1616
Db       1426 CAGAAAGTGCCTGAGGAAGGTTGGACTATGCAAGATGGAACTCCTTGGCCTGGAAACAAC 1485
Qy       1617 GTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGT 1676
Db       1486 GTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTCATAGTGGAGTTCGTGATACG 1545
Qy       1677 GAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAAC 1736
Db       1546 GATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTGAGAAGCGGCCTGGATTTGAT 1605
Qy       1737 CATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAAT 1796
Db       1606 CACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAGTCTCTGCTGTTCTATCAAAC 1665
Qy       1797 GCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAG 1856
Db       1666 GCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCAACAACAGCAAAGCAATTAGA 1725
Qy       1857 GAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTC 1916
Db       1726 GAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGAAAGTTTGTTATGTTCAGTTT 1785
Qy       1917 CCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTT 1976
Db       1786 CCGCAGAGATTTGATGGGATTGATAGACATGATAGATACTCAAACCGTAACGTTGTGTTC 1845
Qy       1977 TTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGA 2036
Db       1846 TTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGACCGATATATGTCGGGACAGGT 1905
Qy       2037 TGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCA 2096
Db       1906 TGTGTGTTTAGAAGACAGGCTCTTTATGGTTTTGATGCACCAAAGAAGAAGAAACCACCA 1965
Qy       2097 TCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGG 2156
Db       1966 GGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGTGTTGTGGGTTGAGAAAGAAG 2025
Qy       2157 AAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAA 2216
Db       2026 AGTAAAACGAAAGCCAAAGATAAGAAAACTAACACTAAAGAGACTTCA------------ 2073
Qy       2217 GAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCT 2276
Db       2074 ------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAGGTGTTATCGTCCCAGTGTCA 2127
Qy       2277 GAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCT 2336
Db       2128 AATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGGAGAAGAAGTTTGGACAATCT 2187
Qy       2337 TCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCT 2396
Db       2188 CCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAGTTCCCCGTAACGCAAGCCCC 2247
Qy       2397 GCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGAC 2456
Db       2248 GCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGAA. 2307
Qy       2457 TGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTC 2516
Db       2308 TGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTGAAGATATCCTGACGGGTTTC 2367
Qy       2517 AAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAA 2576
Db       2368 AAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGCCTAAGCGTGCAGCTTTTAAA 2427
Qy       2577 GGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGG 2636
Db       2428 GGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAGTTCTACGTTGGGCTCTTGGC 2487
```

```
Qy      2637 TCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTG 2696
             |||| | ||||| ||||| ||||  ||||| ||  | |||||||| |||||||| ||| |
Db      2488 TCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGTATGGTTATGGTGGTGGTTTA 2547

Qy      2697 AAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCC 2756
             ||||  ||||| ||||| || ||||||||||||  |||| ||||||||||| ||  ||||
Db      2548 AAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCTATCCTTGGACTTCACTTCCA 2607

Qy      2757 CTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCA 2816
              |  | |  || ||| |  | || ||  | |||||| | ||||| |||||  |||   ||
Db      2608 TTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCACAGGAAAATTCATCGTCCCT 2667

Qy      2817 GAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACG 2876
             |||  |  ||     ||  |  |     ||||||     | || || | |||     |  ||
Db      2668 GAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGTTCATATCCATAGCAGTAACT 2727

Qy      2877 AGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAG 2936
              | |||||  ||||||  ||||  | |||||  |||||  |||||  |||||||| || ||||||
Db      2728 GGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATGATTGGTGGAGAAACGAGCAG 2787

Qy      2937 TTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAG 2996
             ||  |||||  ||  |||||  |  |||||  ||  ||  ||||||  ||||  ||  ||  | |||||
Db      2788 TTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTCTGTTTCAAGGTTTGCTCAAA 2847

Qy      2997 GTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAG 3056
             ||  ||||  ||  |||  |  ||  |  |||||  ||  ||  |||||  |  | |||  ||||  |
Db      2848 GTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAAAAGCAGCAGACGATGGAGCT 2907

Qy      3057 TTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTT 3116
             |||||  |||||  ||  |  |||||  |||||  ||  ||  |||||  |||||  ||  |  ||
Db      2908 TTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGATTCCTCCGACAACACTTCTG 2967

Qy      3117 CTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAG 3176
              |  | |||  |||||||  ||  |  |  ||||||||  ||||  ||  |  ||  ||  |||||
Db      2968 ATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATGCCATTAGCAATGGCTATGAC 3027

Qy      3177 TCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTAT 3236
             ||||||||  ||  ||||||||||  ||  |||||  ||  |||||||  |||||  |||  |  ||
Db      3028 TCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTTGGGTCATTGTTCATTTATAC 3087

Qy      3237 CCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGG 3296
             ||  ||  |||||  ||  ||  ||||  |  ||||  |||  |  ||  ||||||  |  |  ||||||
Db      3088 CCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGCCTACGATTATTGTGGTCTGG 3147

Qy      3297 TCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCG 3356
             ||  |||||  ||  |||||  |||||  |  |||  |  |||||  |  ||  |  ||  ||  |  ||
Db      3148 TCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCAGAGTTAACCCGTTTGTGGCT 3207

Qy      3357 AAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTG 3397
             |   | |  ||  ||  |  | |||   ||||||  ||  ||||
Db      3208 A---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATTG 3245


RESULT 12
AAV06566
ID   AAV06566 standard; cDNA; 3828 BP.
XX
AC   AAV06566;
XX
DT   06-JUL-1998  (first entry)
XX
DE   Arabidopsis cellulose biosynthetic gene clone Ath-A.
XX
KW   Cellulose; cellulose synthase; RSW1 gene; beta-1,4-glucan;
KW   transgenic plant; ss.
XX
OS   Arabidopsis thaliana var. Columbia.
XX
FH   Key             Location/Qualifiers
FT   CDS             239..3490
FT                   /*tag=  a
XX
PN   WO9800549-A1.
XX
PD   08-JAN-1998.
XX
PF   24-JUN-1997;   97WO-AU00402.
XX
PR   27-JUN-1996;   96AU-0000699.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Arioli A,  Betzner AS,  Peng L,  Williamson RE;
XX
DR   WPI; 1998-086974/08.
DR   P-PSDB; AAW33818.
XX
PT   DNA encoding cellulose biosynthetic enzyme - useful for manipulation
PT   of cellulose and beta-1,4-glucan
XX
PS   Claim 12; Page 123-130; 207pp; English.
XX
CC   cDNA clone Ath-A was isolated from an Arabidopsis thaliana cDNA
CC   library using probes produced by PCR with primers (see AAT99632-34)
CC   based on cellulose synthase RSW1 genomic clone 23H12 (see AAV06563)
CC   and EST clone AAT20782 (see AAV06562).  It is closely related to RSW1
CC   cDNA (see AAV06565) and to partial genomic clone 12C4 (see AAV06564),
CC   and encodes a 1084-polypeptide (see AAW33818).  Claimed nucleic acid
CC   molecules (see AAV06562-69) coding for claimed polypeptides (see
CC   AAW33816-20 and AAW46202) involved in cellulose biosynthesis can be
CC   used to manipulate the cellulose and/or beta-glucan content of
CC   transgenic plants.  Expression of nucleic acids in the sense
CC   orientation increases the level of cellulose and reduces the level
CC   of non-crystalline beta-1,4-glucan and starch, providing plants
CC   with modified strength and/or shape and/or fibre properties, or
CC   having increased resistance to stresses or pests.  Antisense,
CC   ribozyme or co-suppression molecules can be used to reduce the
CC   cellulose content of a transgenic plant, e.g. to improve
CC   digestibility or to alter carbon partitioning such that increased
CC   carbon is available for growth, rather than deposited as cellulose.
XX
SQ   Sequence 3828 BP; 1056 A; 721 C; 912 G; 1139 T; 0 other;

  Query Match           36.3%;  Score 1429.4;  DB 19;  Length 3828;
  Best Local Similarity   65.8%;  Pred. No. 0;
  Matches 2172;  Conservative    0;  Mismatches 1066;  Indels   63;  Gaps    4;

Qy       124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
             ||   || |  || ||    ||| |  |  | |   |||| | || ||||| ||||||
Db       219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy       184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
             | |||||  ||||  ||||   |  ||| |    | || |        |    | || |
Db       279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy       244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
              ||  |||||    || || || || || ||  |||   |  |     || | || |
Db       339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy       304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
              ||| || || |||||||| ||||| ||||| || ||  |    |||||  || || || |
Db       399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy       364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
             |  | || || |  ||    || || ||||||||  ||||||||||  |  | || ||
Db       459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy       424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
             | | ||| ||    || || || || || || |  | |  |||        |  |
Db       519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy       484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
                 | ||    ||| |  |  ||  || ||  ||   ||||     | | | | |
Db       579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638
```


Qy 544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
Db 639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT----------CAGATTC 683
Qy 604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
Db 684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743
Qy 664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
Db 744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803
Qy 724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
Db 804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863
Qy 784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
Db 864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923
Qy 844 CGGGGAATGATGGTGGTGGTGATGATGGT----------GACG 876
Db 924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983
Qy 877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
Db 984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043
Qy 937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
Db 1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103
Qy 997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
Db 1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163
Qy 1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
Db 1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223
Qy 1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
Db 1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283
Qy 1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
Db 1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343
Qy 1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
Db 1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403
Qy 1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
Db 1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463
Qy 1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
Db 1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523
Qy 1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
Db 1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583
Qy 1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
Db 1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643
Qy 1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
Db 1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703
Qy 1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
Db 1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763
Qy 1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db 1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823
Qy 1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db 1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883
Qy 1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
Db 1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943
Qy 1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db 1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003
Qy 1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db 2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063
Qy 1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db 2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123
Qy 2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db 2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183
Qy 2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db 2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243
Qy 2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db 2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303
Qy 2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db 2304 AGACTTCA----------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345
Qy 2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db 2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405
Qy 2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db 2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465
Qy 2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db 2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525
Qy 2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db 2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585
Qy 2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db 2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645
Qy 2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db 2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705
Qy 2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db 2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765
Qy 2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736

```
Db       2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825

Qy       2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796
             |||||||||| ||  ||||  |  | |  || ||| |  | || ||  | ||||||  | |
Db       2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885

Qy       2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
             ||||  |||||  |||    ||  |||  |  |  ||     ||   |   |     ||||||     |  |
Db       2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945

Qy       2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
             |  ||  |  |||     |   ||   |  |||||  ||||||   ||||  |  |||||  |||||  ||||
Db       2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005

Qy       2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
             |  ||||||||  ||  ||||||||  |||||  ||  |||||  |   ||||||  ||  ||  ||||||
Db       3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
             ||||  ||  ||   |  |||||  ||   ||||  ||  |||  |  ||  |   |||||  ||  ||  ||||
Db       3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
             |  |   |  |||  ||||   |   |||||  |||||  ||  |   |||||  |||||  ||  ||  ||||
Db       3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
             |  |||||  ||  ||   |  ||   |   |  |||  |||||||  ||   |  |   ||||||||   |||
Db       3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
             |  ||  |   ||  ||  |||||  ||||||||  ||  ||||||||||   ||  |||||  ||   |||
Db       3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
             ||||  |||||  |||  |  ||  ||  ||  |||||  ||   ||  ||||  |  ||||  |||   |   |
Db       3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
             |  ||||||  |   |  ||||||||  |||||  ||  |||||  |||||   |  |||  |  |||||
Db       3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
             |  |||  |  ||  ||   |  ||  |    |  |   ||  ||   |   |  |||    ||||||  ||  |||
Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397
              |
Db       3483 G 3483


RESULT 13
AAA67114
ID   AAA67114 standard; DNA; 3851 BP.
XX
AC   AAA67114;
XX
DT   31-OCT-2000  (first entry)
XX
DE   Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.
XX
KW   Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW   plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW   transgenic plant; ds.
XX
OS   Pinus radiata.
XX
PN   WO200022092-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-NZ00169.
XX
PR   13-OCT-1998;   98US-0170862.
PR   11-AUG-1999;   99US-0148426.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Bloksberg LN;
XX
DR   WPI; 2000-339328/29.
DR   P-PSDB; AAB16307.
XX
PT   New genes encoding proteins involved in a plant polysaccharide
PT   biosynthetic pathway, useful for modulating or altering the
PT   polysaccharide content, composition or structure of the plant  -
XX
PS   Claim 1; Page 71-72; 301pp; English.
XX
CC   The present invention describes isolated polynucleotides (PN) comprising
CC   a sequence selected from one of 835 nucleotide sequences given in
CC   AAA67073 to AAA67907, their (reverse) complements, sequences producing
CC   an Expectation (E) value of 0.01 or less compared to the 835 sequences,
CC   sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC   835 sequences or sequences that are degenerately equivalent or allelic
CC   to the 835 sequences. The polynucleotides are used to modify the
CC   activity of a polypeptide involved in a polysaccharide biosynthetic
CC   pathway in the plant. They are especially used to modulate or alter the
CC   polysaccharide content, composition or structure of the plant. AAB16268
CC   to AAB16340 are proteins encoded by some of the polynucleotide sequence
CC   given in the present invention.
XX
SQ   Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 other;

  Query Match           35.9%;  Score 1411.4;  DB 21;  Length 3851;
  Best Local Similarity 66.3%;  Pred. No. 0;
  Matches 2206;  Conservative    0;  Mismatches 1016;  Indels  105;  Gaps    8;

Qy       123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
             ||  ||  |  |   |  |    | |  |||||  ||  ||||||||   ||||  |||||  ||  ||  |||
Db       132 GGCAGAGAATTGAACTTAACAATGGAAGCCAGCGCCGGCTTGGTTGCCGGTTCTCATAAC 191

Qy       183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAG 242
             |  ||||||  ||||  |||||||       |  ||   ||    |  ||||||||||     |       |
Db       192 AGAAACGAGTTCGTGGTCATCC------ATGGACATGAGGAGCCGAAGCCTTTGAACACG 245

Qy       243 CAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGAC 302
               ||   ||  ||  ||  ||||||||||| |||||  ||||||||  |||   |  |  |  |||  ||
Db       246 TTGAGTGGCCACGTCTGCCAGATTTGTGGCGAGGACGTCGGGCTTAACACAGACGGCGAG 305

Qy       303 CCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAG 362
             |  |||||  ||  ||  ||  |||||||   ||  ||  |||||  |||    |||||  ||  |||||||
Db       306 CTGTTCGTTGCCTGTAATGAGTGCGGGTTTCCTGTCTGTCGGCCGTGCTATGAGTACGAG 365

Qy       363 CGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGC 422
             |  ||  ||  ||  |   |||    |||||  ||||||||  |||||  ||||||||  |   ||  ||
Db       366 AGACGAGAAGGAAATCAGTCGTGCCCGCAGTGCAATACTCGTTACAAGCGTCAAAAAGGG 425

Qy       423 TGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTC 482
             |  |  |||  |||    ||||||||  ||  ||  |  |||  ||||||  |  ||   |  ||  ||
Db       426 AGTCCACGGGTGGAAGGTGACGATGATGAAGAAGACGTTGATGACATAGAACATGAATTT 485

Qy       483 AACTGGGACGGCCA------TGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCAC 536
             ||   |||     ||      ||  ||||    |  |||||  |  ||||||  ||||  |  |
Db       486 AATGTGGAGACTCAGCAAAGAAACAGGCAGCAGATCACCGAGGCGATGCTCCACGGACGC 545

Qy       537 ATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGC------CACAAGCTTTCCAGCTC 590
             ||||||||  |||||  ||     |||     |      |||       |  ||    |    ||||
Db       546 ATGAGCTATGGCCGAGGTCCCGACGACGAAAATTCGCAGATTGCTCATAATCCAGAGCTT 605

Qy       591 AACCCCAATGTTCCACTCCTCACCAACGGGCA---AATGGTGGATGACATCCCACCGGAG 647
             ||   |  ||||  |  ||  |  |||||  ||      |  |||  ||         ||   |
Db       606 CCTCCGCAGATTCCTGTACTTGCAAACGGCCACTCGGTTGTGAGTGGGGAGATTCCAACG 665

Qy       648 CAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCT 707
```

```
Db        666 TCATACTACGCAGACAACCAATTGCTTGCCAACCCTGCAATGCTGAAGCGTGTGCATCCA 725
Qy        708 TATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCT 767
Db        726 AGCTCCGAGCCGGGG------AGTGGAAGGATCATCATGGATCCAAACAGGGATATTGGT 779
Qy        768 GCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAA 827
Db        780 TCTTATGGCTTTGGGAACGTGTCTTGGAAGGAGCGAGGCGATGGTTATAAATCGAAGGAA 839
Qy        828 GAGAGGATGCACCAGACGGGGAATGATGGTGGTGG------------------------- 862
Db        840 AACAAATCAGGCCAGTTGGATATGACGGAAGGGAGATATCAATATAATGGGGGGTTTGCA 899
Qy        863 -----------TGATGATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAA 911
Db        900 CCAAATGAGCCTGAAGATTATATTGATCCCGATATGCCAATGACCGATGAAGCAAGGCAG 959
Qy        912 CAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATC 971
Db        960 CCACTGTCCCGAAAAGTGCCAATTCCTTCAAGCAAAATAAATCCATACCGAATGGTCATT 1019
Qy        972 ATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAAT 1031
Db       1020 GTAATTCGACTGATAGTGCTGGGTATTTTTCTCCGCTATCGTCTCCTGAATCCAGTGAAG 1079
Qy       1032 GATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGG 1091
Db       1080 AATGCATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGGTTTGCCTTGTCATGG 1139
Qy       1092 ATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTG 1151
Db       1140 ATTCTTGATCAGTTTCCCAAGTGGTTGCCTATCAGTCGTGAAACGTATCCTGATCGACTG 1199
Qy       1152 TCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTC 1211
Db       1200 TCATTAAGGTACGAACGAGAAGGCGAACCATCAATGCTTGCACCTGTTGACCTCTTTGTG 1259
Qy       1212 AGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATC 1271
Db       1260 AGTACTGTAGATCCACTGAAGGAGCCTCCTTTGGTTACTGCCAATACAGTATTATCAATC 1319
Qy       1272 CTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCA 1331
Db       1320 CTTTCAGTAGACTACCCTGTAGACAATGTGTCCTGTTATGTCTCTGATGACGGAGCGTCG 1379
Qy       1332 ATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTC 1391
Db       1380 ATGCTTACTTTTGAATCTCTCTCTGAGACCTCAGAATTTGCCAGAAAATGGGTACCATTC 1439
Qy       1392 TGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGAC 1451
Db       1440 TGCAAGAAATTCGACATTGAGCCTCGCGCTCCCGAAATCTATTTCTCTCAGAAAATTGAC 1499
Qy       1452 TACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAG 1511
Db       1500 TATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAGAGCGCCGTGCCATGAAGAGAGAA 1559
Qy       1512 TATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAA 1571
Db       1560 TATGAAGAATTCAAGGTGCGCATCAATCGGTTGGTTGCAAAGGCCTCTAAAGTGCCCAAG 1619
Qy       1572 GAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCT 1631
Db       1620 GAAGGATGGACAATGCAAGACGGTACGCCTTGGCCTGGTAATAATACCCGTGACCATCCT 1679
Qy       1632 GGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTG 1691
Db       1680 GGTATGATCCAAGTGTTCTTGGGTCACAGTGGCGGCCTCGATACAGAAGGCAATGAGCTT 1739
Qy       1692 CCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCT 1751
```

```
Db       1740 CCTCGGCTAGTATATGTTTCTCGTGAGAAGAGACCTGGTTTCCAGCATCACAAGAAGGCC 1799
Qy       1752 GGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTA 1811
Db       1800 GGTGCCATGAATGCTTTGGTTCGGGTTTCTGCTGTGCTCACCAATGCTCCATTTATGCTG 1859
Qy       1812 AACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTT 1871
Db       1860 AATCTGGATTGTGATCACTACATTAACAATAGCAAGGCAATCAGGGAAGGCATGTGCTTT 1919
Qy       1872 ATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGAT 1931
Db       1920 ATGATGGATCCTCAGGTTGGGAGAAAAGTCTGTTATGTCCAATTCCCTCAGAGATTCGAT 1979
Qy       1932 GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATG 1991
Db       1980 GGTATTGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTCTTTGATATCAACATG 2039
Qy       1992 AAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGG 2051
Db       2040 AAAGGTCTGGATGGAATTCAAGGGCCTGTATATGTGGGAACTGGATGCATGTTCAGAAGA 2099
Qy       2052 CAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAAC 2111
Db       2100 CAAGCTCTATATGGGTATGGGCCTCCCAAAGGCCCAAAAC------------------- 2139
Qy       2112 TGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACT 2171
Db       2140 ----GTCCCAAGATGGTGACCTGTGATTGTCTCCCTTGTTGCGGTCCTCGTAAGAAGTCT 2195
Qy       2172 ACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCC 2231
Db       2196 CCGAAGAAAAATAGTAGCAAGAAAAG------------------TGCAGGAATCCCAGCT 2237
Qy       2232 CCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCC 2291
Db       2238 CCCGCCTACAATCTGGACGGGATCGAGGAAGGAGTAGAAGGTTATGATGACGAAAGAGCA 2297
Qy       2292 GGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACA 2351
Db       2298 TTGTTGATGAGCCAACTAGACTTCGAGAAGAAGTTTGGCCAGTCTTCAGCTTTTGTTCAA 2357
Qy       2352 TCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAA 2411
Db       2358 TCCACTCTGATGGAGAATGGTGGTGTTCCGCAAACAGCAAATCCAGCTGAATTGTTGAAG 2417
Qy       2412 GAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATT 2471
Db       2418 GAGGCTATTCATGTCATCAGCTGTGGATATGAAGACAAAACGGAATGGGGAAAAGAGCTT 2477
Qy       2472 GGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCAT 2531
Db       2478 GGATGGATCTATGGATCAGTCACAGAGGACATTCTGACTGGATTCAAGATGCACACTCGA 2537
Qy       2532 GGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTG 2591
Db       2538 GGCTGGCGGTCCATTTACTGTATGCCCAAACGAGCAGCATTCAAAGGGTCTGCTCCAATC 2597
Qy       2592 AATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651
Db       2598 AATCTATCAGACCGTTTGAACCAGGTGTTGCGTTGGGCTTTGGGATCAGTAGAAATTTTC 2657
Qy       2652 TTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGA 2711
Db       2658 ATGAGCAGACATTGCCCAATCTGGTATGGCTATGGGGGAGGTCTGAAATGGCTTGAAAGA 2717
Qy       2712 TTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGT 2771
Db       2718 TTTGCCTATATCAACACCATTGTCTATCCATTCACCTCTCTTCCACTCATTGCCTATTGC 2777
Qy       2772 ACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTT 2831
Db       2778 ACACTTCCAGCCGTCAGTTTGCTCACTGGCAAATTTGTGATCCCTCAGATCAGTACTTTT 2837
```

```
Qy       2832 GCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATG 2891
             || || || |  || ||  | ||||| ||||  |||||||| ||  | || ||  |||||
Db       2838 GCAAGTCTATTTTTAATAGCTCTTTTCATCTCAATTTTTGCCACTGGTATTCTGAAAATG 2897

Qy       2892 AGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGA 2951
             || |||||||| ||  | ||||| || |||||| | ||||| |||||||||| ||||||
Db       2898 AGGTGGAGTGGAGTGAGCATTGAAGAATGGTGGCGAAATGAACAGTTCTGGGTTATTGGA 2957

Qy       2952 GGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTT 3011
             || || ||  |||| | ||||| ||  | || || || ||||||||  | || ||  ||
Db       2958 GGGGTTTCTGCACATTTTTTTGCAGTTATTCAAGGTCTGCTCAAGGTACTGGCAGGCATT 3017

Qy       3012 GATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
             |||||||  ||||| || ||  | ||||    |||||| |  |||||    || || |||
Db       3018 GATACAAATTTCACAGTCACTGCCAAGGCATCAGATGACGGTGAGTTTGGGGAACTGTAT 3077

Qy       3072 ACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATT 3131
             ||||||||||||| ||  |  | || |||||||| ||| ||||| |  | ||| |  |
Db       3078 GCATTCAAATGGACCACACTCCTCATTCCTCCTACAACCCTGCTTGTCATCAACCTTGTG 3137

Qy       3132 GGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTC 3191
             || ||||| | |||||| || |||| ||||| || |||| | |||||||||| || |||
Db       3138 GGGGTGGTTGTTGGCGTAGCAGATGCAATCAACAATGGATTTCAGTCATGGGGTCCTCTC 3197

Qy       3192 TTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGT 3251
             || || ||||| |||||||||| ||||| ||||| || || ||||| || ||||| |||
Db       3198 TTGGGTAAGCTTTTCTTTGCATTCTGGGTCATTGTGCACCTGTATCCTTTCCTCAAGGGT 3257

Qy       3252 TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCT 3311
               | || ||||| ||| | ||||| || || ||  | | ||||| ||||||||||
Db       3258 CTCATGGGCAGGCAGAACCGAACACCCACCATCGTGGTTATTTGGTCAATTCTGCTGGCA 3317

Qy       3312 TCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCG 3371
             ||  | ||||| ||  | ||||||  | |||||||||||| |    |||| | | || ||
Db       3318 TCTGTTTTCTCTCTTTTCTGGGTAAGAATTGATCCTTTCTTGAGTAAGGTTAAAGGCCCA 3377

Qy       3372 CTTCTTGAGGAGTGTGGTTTGGATTGC 3398
                | | |  | |||||  |  | |||
Db       3378 GATACTAAACAATGTGGCATCAACTGC 3404
```

RESULT 14
AAZ99500
ID AAZ99500 standard; DNA; 3725 BP.
XX
AC AAZ99500;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key   Location/Qualifiers
FT CDS   179..3398
FT   /*tag= a
FT   /product= "cellulose synthase"
FT   /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84110.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 1; Page 105-110; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

Query Match 35.0%; Score 1378.4; DB 21; Length 3725;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;

```
Qy        119 CGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCA 178
              ||| |   || |  ||   |  | ||||| ||| | |   ||| ||||||| ||||| ||
Db        154 CGGTGTGTGTAGGAGGAGCGCGGAGATGGCGGCCAACAAGGGGATGGTGGCAGGCTCTCA 213

Qy        179 CAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA------TCCCGGGCCGAAGCC 232
              |||||||||||||| ||||| | ||||||| ||||||||      |  |  | | |||||
Db        214 CAACCGCAACGAGTTCGTCATGATCCGCCACGACGGCGACGCGCCTGTCCCGGCTAAGCC 273

Qy        233 GCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCC 292
                ||  |     ||| |||||||| |||||||||| ||||||   || ||| || |  |
Db        274 CACGAAGAGTGCGAATGGGCAGGTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTCAGC 333

Qy        293 CGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTA 352
              |   || ||   ||| || || ||||| ||||| |||||||| ||||||||    |||||
Db        334 CACTGGTGATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCTTGCTA 393

Qy        353 CGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCG 412
              || |||||||||  ||| || |  ||    |||||| |||||||||||| |||||||| |
Db        394 TGAGTACGAGCGCAAGGAAGGGAACCAATGCTGCCCTCAGTGCAAGACTAGATACAAGAG 453

Qy        413 CCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGA 472
              |  || ||  |||  || ||    ||||| || |||||||| |  || ||||||||||
Db        454 ACAGAAAGGTAGCCCTCGAGTTCATGGTGATGATGAGGAGGAAGATGTTGATGACCTGGA 513

Qy        473 CAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGG 532
              ||| || |||||||   |                    |   ||     | ||
Db        514 CAATGAATTCAACTATAA----------------------GCAAGGCAATGGGAAGGG 549

Qy        533 CCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAA 592
              ||    | |  |    |  |||| |||| || ||    |  ||   |||  |||   |
Db        550 CCCAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCC 609

Qy        593 CCCCAAT---GTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCA 649
                | | ||    |||||| ||| || |  || |||  | |   ||        || ||
```

```
Db        610 ACACCATCGGATTCCACGCCTTACAAGTGGACAACAGATATCTGGAGAGATCCCTGATGC 669
Qy        650 GCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTA 709
Db        670 ATCCCCTGACCGTCATTCTATCCGCAGT-----------------CCAACATCGAGCTA 711
Qy        710 TGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGC 769
Db        712 TGTTGATCCAAGCGTTCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTC 771
Qy        770 ATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA 829
Db        772 CTATGGGCTTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831
Qy        830 GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG---------------- 873
Db        832 CAAAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGGAC 891
Qy        874 --------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAA 925
Db        892 TGGCTCAAATGGAGAAGATATGCAAATGGTTGATGATGCACGCCTACCTTTGAGCCGCAT 951
Qy        926 AATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGT 985
Db        952 TGTGCCAATTTCCTCAAACCAGCTCAACCTTTACCGGATAGTAATCATTCTCCGTCTTAT 1011
Qy        986 GGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTT 1045
Db       1012 CATCCTGTGCTTCTTCTTCCAATATCGTATCAGTCATCCAGTGCGTAATGCTTATGGATT 1071
Qy       1046 GTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATT 1105
Db       1072 GTGGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGCCTTGTCCTGGCTTCTAGATCAGTT 1131
Qy       1106 CCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGA 1165
Db       1132 CCCAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTGAGGTATGA 1191
Qy       1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225
Db       1192 TAGAGAGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTTGTCAGTACAGTGGATCC 1251
Qy       1226 CTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTA 1285
Db       1252 ATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTGCTGTGGATTA 1311
Qy       1286 TCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGA 1345
Db       1312 CCCTGTTGACAAAGTGTCATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGA 1371
Qy       1346 AGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAA 1405
Db       1372 GTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGAAGCACAA 1431
Qy       1406 TATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAA 1465
Db       1432 TATTGAACCAAGAGCTCCAGAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAA 1491
Qy       1466 GGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAA 1525
Db       1492 AATTCAACCTTCATTTGTTAAGGAAAGACGAGCAATGAAGAGAGAGTATGAAGAATTCAA 1551
Qy       1526 GGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAAT 1585
Db       1552 AATAAGAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCAT 1611
Qy       1586 GCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGT 1645
Db       1612 GGCTGATGGAACTGCTTGGCCTGGGAATAACCCTAGGGACCATCCTGGCATGATTCAGGT 1671
Qy       1646 CTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTA 1705
Db       1672 GTTCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAAATGAATTACCACGTCTTGTCTA 1731

Qy       1706 TGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGC 1765
Db       1732 TGTCTCTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAATGAATGC 1791
Qy       1766 ATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGA 1825
Db       1792 ACTGATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGTGA 1851
Qy       1826 TCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTT 1885
Db       1852 CCATTACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCAGC 1911
Qy       1886 ACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCA 1945
Db       1912 TCTAGGAAGGAAAACTTGTTATGTACAATTTCCACAAAGATTTGATGGCATTGACTTGCA 1971
Qy       1946 TGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGG 2005
Db       1972 CGATCGATATGCTAATAGGAACATAGTCTTCTTTGATATCAACATGAAAGGTCTAGATGG 2031
Qy       2006 TATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGG 2065
Db       2032 CATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGG 2091
Qy       2066 TTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTG 2125
Db       2092 ATATGAT-----------------CCTGTTTTGACTGAAGCTGATCTGGAACCTAACATT 2134
Qy       2126 GTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAAC 2185
Db       2135 GTTGTTAAG----AGCTGCTGTGGTAGAAGGAAGAGAAAGAACAAGAGTTATATGGATAG 2190
Qy       2186 AGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCT 2245
Db       2191 TCAAAGCCGTA---------TTATGAAGAGAACAGAATCTTCAGCTCCCATCTTTAACAT 2241
Qy       2246 TGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCA 2305
Db       2242 GGAAGACATCGAGGAGGGTATTGAAGGTTATGAGGATGAAAGGTCAGTGCTTATGTCCCA 2301
Qy       2306 ACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGA 2365
Db       2302 GAGGAAATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTCATTGCATCCACCTTTATGAC 2361
Qy       2366 GAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGT 2425
Db       2362 TCAAGGTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTGAAGGAAGCTATCCATGT 2421
Qy       2426 CATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGG 2485
Db       2422 TATCAGCTGTGGGTACGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGATCTATGG 2481
Qy       2486 ATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAAT 2545
Db       2482 TTCAGTTACAGAGGATATTCTGACTGGGTTTAAAATGCATGCAAGAGGCTGGCAATCAAT 2541
Qy       2546 TTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCG 2605
Db       2542 CTACTGCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATCG 2601
Qy       2606 TCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTG 2665
Db       2602 TCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGCAGACATTG 2661
Qy       2666 CCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAA 2725
Db       2662 TCCTATATGGTATGGCTACAATGGGCGATTGAAGCTTTTGGAGAGGCTGGCTTACATTAA 2721
Qy       2726 CTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCAT 2785
Db       2722 CACCATTGTTTATCCAATCACATCTGTTCCGCTTATCGCCTATTGTGTGCTTCCTGCTAT 2781
```

```
Qy       2786 CTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTT 2845
Db       2782 CTGTCTTCTTACCAATAAATTTATCATTCCTGAGATTAGTAATTATGCTGGAATGTTCTT 2841

Qy       2846 CATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGT 2905
Db       2842 CATTCTTCTTTTTGCCTCCATTTTCGCAACTGGTATATTGGAGCTCAGATGGAGTGGTGT 2901

Qy       2906 TGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACA 2965
Db       2902 TGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCTCTGCCCA 2961

Qy       2966 CCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCAC 3025
Db       2962 TCTCTTCGCGGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCAACTTCAC 3021

Qy       3026 CGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATG 3082
Db       3022 AGTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTG 3081

Qy       3083 GACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGC 3142
Db       3082 GACCAGTTTGCTCATCCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAATGGTGGC 3141

Qy       3143 TGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCT 3202
Db       3142 AGGAATTTCGTATGCCATTAACAGCGGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCT 3201

Qy       3203 ATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAG 3262
Db       3202 GTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCCTCAAGGGTCTCATGGGCAG 3261

Qy       3263 GCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTC 3322
Db       3262 GCAGAACCGCACGCCAACAATCGTCATCGTTTGGTCCATCCTCCTTGCGTCTATCTTCTC 3321

Qy       3323 GCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352
Db       3322 CTTGCTGTGGGTGAAGATCGATCCTTTCAT 3351
```

```
RESULT 15
AAZ99506
ID   AAZ99506 standard; DNA; 3725 BP.
XX
AC   AAZ99506;
XX
DT   03-JUL-2000  (first entry)
XX
DE   DNA encoding a maize cellulose synthase.
XX
KW   Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW   transgenic plant; plant breeding marker; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             179..3397
FT                   /*tag=  a
FT                   /product= "cellulose synthase"
FT                   /note= "no termination codon given"
XX
PN   WO200009706-A2.
XX
PD   24-FEB-2000.
XX
PF   16-AUG-1999;   99WO-US18760.
XX
PR   17-AUG-1998;   98US-0096822.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Dhugga KS,  Helentjaris TG,  Bowen BA,  Wang X;
XX
DR   WPI; 2000-224343/19.
DR   P-PSDB; AAY84112.
XX
PT   New genes which encode maize cellulose synthase polypeptides in plants
PT   useful for modulating the expression of cellulose synthase in plants
PT   and to produce transgenic plants expressing the novel protein  -
XX
PS   Claim 1; Page 121-126; 119pp; English.
XX
CC   The present sequence encodes a maize cellulose synthase polypeptide.
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

  Query Match             35.0%;  Score 1378.4;  DB 21;  Length 3725;
  Best Local Similarity   66.4%;  Pred. No. 0;
  Matches 2171;  Conservative    0;  Mismatches  991;  Indels  108;  Gaps    9;

Qy       119 CGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCA 178
Db       154 CGGTGTGTGTAGGAGGAGCGCGGAGATGGCGGCCAACAAGGGGATGGTGGCAGGCTCTCA 213

Qy       179 CAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA------TCCCGGGCCGAAGCC 232
Db       214 CAACCGCAACGAGTTCGTCATGATCCGCCACGACGGCGACGCGCCTGTCCCGGCTAAGCC 273

Qy       233 GCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCC 292
Db       274 CACGAAGAGTGCGAATGGGCAGGTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTCAGC 333

Qy       293 CGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTA 352
Db       334 CACTGGTGATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCTTGCTA 393

Qy       353 CGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCG 412
Db       394 TGAGTACGAGCGCAAGGAAGGGAACCAATGCTGCCCTCAGTGCAAGACTAGATACAAGAG 453

Qy       413 CCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGA 472
Db       454 ACAGAAAGGTAGCCCTCGAGTTCATGGTGATGATGAGGAGGAAGATGTTGATGACCTGGA 513

Qy       473 CAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGG 532
Db       514 CAATGAATTCAACTATAA----------------------GCAAGGCAATGGGAAGGG 549

Qy       533 CCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAA 592
Db       550 CCCAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCC 609

Qy       593 CCCCAAT---GTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCA 649
Db       610 ACACCATCGGATTCCACGCCTTACAAGTGGACAACAGATATCTGGAGAGATCCCTGATGC 669
```

```
Qy       650 GCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTA 709
Db       670 ATCCCCTGACCGTCATTCTATCCGCAGT------------------CCAACATCGAGCTA 711

Qy       710 TGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGC 769
Db       712 TGTTGATCCAAGCGTTCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTC 771

Qy       770 ATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA 829
Db       772 CTATGGGCTTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831

Qy       830 GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG---------------- 873
Db       832 CAAAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGGAC 891

Qy       874 --------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAA 925
Db       892 TGGCTCAAATGGAGAAGATATGCAAATGGTTGATGATGCACGCCTACCTTTGAGCCGCAT 951

Qy       926 AATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGT 985
Db       952 TGTGCCAATTTCCTCAAACCAGCTCAACCTTTACCGGATAGTAATCATTCTCCGTCTTAT 1011

Qy       986 GGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTT 1045
Db      1012 CATCCTGTGCTTCTTCTTCCAATATCGTATCAGTCATCCAGTGCGTAATGCTTATGGATT 1071

Qy      1046 GTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATT 1105
Db      1072 GTGGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGCCTTGTCCTGGCTTCTAGATCAGTT 1131

Qy      1106 CCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGA 1165
Db      1132 CCCAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTGAGGTATGA 1191

Qy      1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225
Db      1192 TAGAGAGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTTGTCAGTACAGTGGATCC 1251

Qy      1226 CTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTA 1285
Db      1252 ATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTGCTGTGGATTA 1311

Qy      1286 TCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGA 1345
Db      1312 CCCTGTTGACAAAGTGTCATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGA 1371

Qy      1346 AGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAA 1405
Db      1372 GTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGAAGCACAA 1431

Qy      1406 TATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAA 1465
Db      1432 TATTGAACCAAGAGCTCCAGAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAA 1491

Qy      1466 GGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAA 1525
Db      1492 AATTCAACCTTCATTTGTTAAGGAAAGACGAGCAATGAAGAGAGAGTATGAAGAATTCAA 1551

Qy      1526 GGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAAT 1585
Db      1552 AATAAGAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCAT 1611

Qy      1586 GCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGT 1645
Db      1612 GGCTGATGGAACTGCTTGGCCTGGGAATAACCCTAGGGACCATCCTGGCATGATTCAGGT 1671

Qy      1646 CTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTA 1705
Db      1672 GTTCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAAATGAATTACCACGTCTTGTCTA 1731

Qy      1706 TGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGC 1765
Db      1732 TGTCTCTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAATGAATGC 1791

Qy      1766 ATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGA 1825
Db      1792 ACTGATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGTGA 1851

Qy      1826 TCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTT 1885
Db      1852 CCATTACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCAGC 1911

Qy      1886 ACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCA 1945
Db      1912 TCTAGGAAGGAAAACTTGTTATGTACAATTTCCACAAAGATTTGATGGCATTGACTTGCA 1971

Qy      1946 TGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGG 2005
Db      1972 CGATCGATATGCTAATAGGAACATAGTCTTCTTTGATATCAACATGAAAGGTCTAGATGG 2031

Qy      2006 TATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGG 2065
Db      2032 CATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGG 2091

Qy      2066 TTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTG 2125
Db      2092 ATATGAT-----------------CCTGTTTTGACTGAAGCTGATCTGGAACCTAACATT 2134

Qy      2126 GTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAAC 2185
Db      2135 GTTGTTAAG----AGCTGCTGTGGTAGAAGGAAGAGAAAGAACAAGAGTTATATGGATAG 2190

Qy      2186 AGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCT 2245
Db      2191 TCAAAGCCGTA---------TTATGAAGAGAACAGAATCTTCAGCTCCCATCTTTAACAT 2241

Qy      2246 TGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCA 2305
Db      2242 GGAAGACATCGAGGAGGGTATTGAAGGTTATGAGGATGAAAGGTCAGTGCTTATGTCCCA 2301

Qy      2306 ACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGA 2365
Db      2302 GAGGAAATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTCATTGCATCCACCTTTATGAC 2361

Qy      2366 GAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGT 2425
Db      2362 TCAAGGTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTGAAGGAAGCTATCCATGT 2421

Qy      2426 CATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGG 2485
Db      2422 TATCAGCTGTGGGTACGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGATCTATGG 2481

Qy      2486 ATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAAT 2545
Db      2482 TTCAGTTACAGAGGATATTCTGACTGGGTTTAAAATGCATGCAAGAGGCTGGCAATCAAT 2541

Qy      2546 TTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCG 2605
Db      2542 CTACTGCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATCG 2601

Qy      2606 TCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTG 2665
Db      2602 TCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGCAGACATTG 2661

Qy      2666 CCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAA 2725
Db      2662 TCCTATATGGTATGGCTACAATGGGCGATTGAAGCTTTTGGAGAGGCTGGCTTACATTAA 2721

Qy      2726 CTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCAT 2785
Db      2722 CACCATTGTTTATCCAATCACATCTGTTCCGCTTATCGCCTATTGTGTGCTTCCTGCTAT 2781

Qy      2786 CTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTT 2845
```

```
Db       2782 CTGTCTTCTTACCAATAAATTTATCATTCCTGAGATTAGTAATTATGCTGGAATGTTCTT 2841

Qy       2846 CATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGT 2905

Db       2842 CATTCTTCTTTTTGCCTCCATTTTCGCAACTGGTATATTGGAGCTCAGATGGAGTGGTGT 2901

Qy       2906 TGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACA 2965

Db       2902 TGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCTCTGCCCA 2961

Qy       2966 CCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCAC 3025

Db       2962 TCTCTTCGCGGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCAACTTCAC 3021

Qy       3026 CGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATG 3082

Db       3022 AGTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTG 3081

Qy       3083 GACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGC 3142

Db       3082 GACCAGTTTGCTCATCCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAATGGTGGC 3141

Qy       3143 TGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCT 3202

Db       3142 AGGAATTTCGTATGCCATTAACAGCGGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCT 3201

Qy       3203 ATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAG 3262

Db       3202 GTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCCTCAAGGGTCTCATGGGCAG 3261

Qy       3263 GCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTC 3322

Db       3262 GCAGAACCGCACGCCAACAATCGTCATCGTTTGGTCCATCCTCCTTGCGTCTATCTTCTC 3321

Qy       3323 GCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352

Db       3322 CTTGCTGTGGGTGAAGATCGATCCTTTCAT 3351

Search completed: December 15, 2003, 18:48:39
Job time : 979 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 22:42:15 ; Search time 1271 Seconds
(without alignments)
10292.436 Million cell updates/sec

Title: US-09-720-383C-9
Perfect score: 3936
Sequence: 1 cttctccctcgtcggtgcgg.....ctgagttggttttgtgaaaa 3936

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3936	100.0	3936	10	US-09-900-237-9	Sequence 9, Appli
2	3936	100.0	3969	13	US-10-209-059-49	Sequence 49, Appl
3	3936	100.0	3969	13	US-10-160-719-17	Sequence 17, Appl
4	3936	100.0	3969	13	US-10-160-719-37	Sequence 37, Appl
5	2592.6	65.9	3776	10	US-09-900-237-3	Sequence 3, Appli
6	2546.8	64.7	3568	13	US-10-209-059-41	Sequence 41, Appl
7	2546.8	64.7	3568	13	US-10-160-719-1	Sequence 1, Appli
8	2241	56.9	3786	10	US-09-900-237-7	Sequence 7, Appli
9	2241	56.9	3813	13	US-10-209-059-17	Sequence 17, Appl
10	2241	56.9	3813	13	US-10-160-719-25	Sequence 25, Appl
11	2241	56.9	3813	13	US-10-160-719-45	Sequence 45, Appl
12	1509	38.3	3255	10	US-09-938-842A-2530	Sequence 2530, Ap
13	1431.8	36.4	3255	10	US-09-938-842A-559	Sequence 559, App
14	1429.4	36.3	3828	15	US-10-229-193-7	Sequence 7, Appli
15	1378.4	35.0	3725	13	US-10-209-059-45	Sequence 45, Appl
16	1378.4	35.0	3725	13	US-10-160-719-13	Sequence 13, Appl
17	1378.4	35.0	3725	13	US-10-160-719-21	Sequence 21, Appl
18	1378.4	35.0	3725	13	US-10-160-719-41	Sequence 41, Appl
19	1365	34.7	3753	13	US-10-160-719-33	Sequence 33, Appl
20	1365	34.7	3753	13	US-10-160-719-53	Sequence 53, Appl
21	1365	34.7	3780	13	US-10-209-059-1	Sequence 1, Appli
22	1365	34.7	3780	13	US-10-160-719-9	Sequence 9, Appli
23	1331.8	33.8	2890	10	US-09-900-237-17	Sequence 17, Appl
24	1330.8	33.8	3673	15	US-10-229-193-11	Sequence 11, Appl
25	1326.6	33.7	3603	15	US-10-229-193-5	Sequence 5, Appli
26	1322.4	33.6	3704	13	US-10-209-059-13	Sequence 13, Appl
27	1322.4	33.6	3704	13	US-10-160-719-57	Sequence 57, Appl
28	1288.2	32.7	3746	13	US-10-160-719-29	Sequence 29, Appl
29	1288.2	32.7	3746	13	US-10-160-719-49	Sequence 49, Appl
30	1288.2	32.7	3773	13	US-10-209-059-9	Sequence 9, Appli
31	1288.2	32.7	3773	13	US-10-160-719-5	Sequence 5, Appli
32	1287.2	32.7	3563	10	US-09-900-237-25	Sequence 25, Appl
33	1279.4	32.5	3799	13	US-10-267-459-5	Sequence 5, Appli
34	1279.4	32.5	3799	13	US-10-209-059-21	Sequence 21, Appl
35	1271	32.3	3626	10	US-09-900-237-29	Sequence 29, Appl
36	1269.2	32.2	3614	15	US-10-229-193-9	Sequence 9, Appli
37	1254.2	31.9	3517	10	US-09-900-237-13	Sequence 13, Appl
38	1218	30.9	2663	10	US-09-900-237-23	Sequence 23, Appl
39	1170	29.7	2830	13	US-10-267-459-1	Sequence 1, Appli
40	1170	29.7	2830	13	US-10-209-059-5	Sequence 5, Appli
41	1138.8	28.9	3231	13	US-10-209-059-27	Sequence 27, Appl
42	1134.8	28.8	3328	10	US-09-838-539-1	Sequence 1, Appli
43	1016.6	25.8	3028	13	US-10-209-059-29	Sequence 29, Appl
44	1004	25.5	3470	13	US-10-209-059-25	Sequence 25, Appl
45	884.6	22.5	2248	15	US-10-229-193-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-900-237-9
; Sequence 9, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 3936
; TYPE: DNA
; ORGANISM: Zea mays
US-09-900-237-9

Query Match 100.0%; Score 3936; DB 10; Length 3936;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
Db        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
Db        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
Db        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
Db        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Db       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Db       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Db       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
```

```
Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Db       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Db       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Db       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Db       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Db       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Db       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Db       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Db       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Db       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Db       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Qy       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Qy       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Qy       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Qy       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Qy       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Qy       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Qy       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Qy       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Db       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
```

```
Qy       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400

Qy       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460

Qy       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520

Qy       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580

Qy       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640

Qy       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700

Qy       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760

Qy       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820

Qy       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880

Qy       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940

Qy       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000

Qy       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060

Qy       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120

Qy       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180

Qy       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240

Qy       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300

Qy       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360

Qy       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420

Qy       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480

Qy       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540

Qy       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600

Qy       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660

Qy       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720

Qy       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780

Qy       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840

Qy       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900

Qy       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             ||||||||||||||||||||||||||||||||||||
Db       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 2
US-10-209-059-49
; Sequence 49, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-49

Query Match 100.0%; Score 3936; DB 13; Length 3969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
```

```
Qy       121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
Db       121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy       181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
Db       181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy       241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
Db       241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy       301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
Db       301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy       361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
Db       361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy       421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Db       421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480

Qy       481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Db       481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540

Qy       541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Db       541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600

Qy       601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Db       601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660

Qy       661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Db       661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720

Qy       721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Db       721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780

Qy       781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Db       781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840

Qy       841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Db       841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900

Qy       901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Db       901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960

Qy       961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Db       961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020

Qy      1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Db      1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080

Qy      1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
Db      1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140

Qy      1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Db      1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
```

```
Qy      1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Db      1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260

Qy      1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Db      1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320

Qy      1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Db      1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380

Qy      1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Db      1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440

Qy      1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Db      1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500

Qy      1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Db      1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560

Qy      1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Db      1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620

Qy      1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Db      1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680

Qy      1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Db      1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740

Qy      1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Db      1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800

Qy      1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Db      1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860

Qy      1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db      1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920

Qy      1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db      1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980

Qy      1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db      1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040

Qy      2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db      2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100

Qy      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160

Qy      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220

Qy      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280

Qy      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
```

```
Db      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Qy      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Db      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Qy      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Db      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Qy      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Db      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Qy      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Db      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Qy      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Db      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Qy      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Db      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Qy      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Db      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Qy      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Db      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Qy      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Db      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Qy      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Db      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Qy      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Db      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Qy      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Db      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Qy      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Db      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Qy      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Db      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Qy      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Db      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Qy      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Db      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Qy      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
Db      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
Qy      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
Db      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
Qy      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
Db      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
Qy      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
Db      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
Qy      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
Db      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
Qy      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
Db      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
Qy      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
Db      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
Qy      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
Db      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
Qy      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
Db      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
Qy      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
Db      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
Qy      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
Db      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 3
US-10-160-719-17
; Sequence 17, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719-17

Query Match 100.0%; Score 3936; DB 13; Length 3969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
```

```
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
Db        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
Db        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
Db        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
Db        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
Db        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480

Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540

Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600

Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660

Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720

Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780

Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840

Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900

Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960

Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020

Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080

Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
```

```
Db       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140

Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
Db       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200

Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
Db       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260

Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
Db       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320

Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
Db       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380

Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
Db       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440

Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
Db       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500

Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
Db       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560

Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
Db       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620

Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
Db       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680

Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
Db       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740

Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
Db       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800

Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
Db       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860

Qy       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
Db       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920

Qy       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
Db       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980

Qy       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
Db       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040

Qy       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
Db       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100

Qy       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Db       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160

Qy       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db       2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
```

```
Qy       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280

Qy       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340

Qy       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400

Qy       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460

Qy       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520

Qy       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580

Qy       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640

Qy       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700

Qy       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760

Qy       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820

Qy       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880

Qy       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940

Qy       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000

Qy       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060

Qy       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120

Qy       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180

Qy       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240

Qy       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300

Qy       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360

Qy       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420

Qy       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCATATTCAGTGT 3480

Qy       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540

Qy       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600

Qy       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660

Qy       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720

Qy       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780

Qy       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840

Qy       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900

Qy       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
             ||||||||||||||||||||||||||||||||||||
Db       3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 4
US-10-160-719-37
; Sequence 37, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719-37

Query Match 100.0%; Score 3936; DB 13; Length 3969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTCTCCCTCGTCGGTGCGGCGTGGCGCGGCTCGGCGTTCGGTGAGAAACCACTCGGGGG 60

Qy         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGGATCTGCTGCTAGAGTGAGAGGAGCTACGGTCAGTATCCTCTGCCTTCGTCGGCG 120

Qy        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACA 180

Qy        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGG 240

Qy        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGG 300

Qy        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACG 360

Qy        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGG 420

Qy        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGT 480

Qy        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGA 540

Qy        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATG 600

Qy        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGG 660

Qy        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCA 720

Qy        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATG 780

Qy        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACC 840

Qy        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATG 900

Qy        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATA 960

Qy        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGC 1020

Qy       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTG 1080

Qy       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACC 1140

Qy       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTG 1200

Qy       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTG 1260

Qy       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATG 1320

Qy       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAAT 1380

Qy       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAAC 1440

Qy       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAA 1500

Qy       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGA 1560

Qy       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTC 1620

Qy       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGG 1680

Qy       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATC 1740

Qy       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800

Qy       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860

Qy       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTC 1920

Qy       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTG 1980

Qy       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040

Qy       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAA 2100

Qy       2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
```

```
Db      2101 GGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGC 2160
Qy      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Db      2161 AAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAG 2220
Qy      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Db      2221 AGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGA 2280
Qy      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Db      2281 ATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTG 2340
Qy      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Db      2341 TTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTT 2400
Qy      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Db      2401 CTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGG 2460
Qy      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Db      2461 GAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGA 2520
Qy      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Db      2521 TGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTT 2580
Qy      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Db      2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTA 2640
Qy      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Db      2641 TTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAAT 2700
Qy      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Db      2701 TTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCT 2760
Qy      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Db      2761 TGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGC 2820
Qy      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Db      2821 TGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCA 2880
Qy      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Db      2881 TCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCT 2940
Qy      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Db      2941 GGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000
Qy      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Db      3001 TAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCT 3060
Qy      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Db      3061 CAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTAT 3120
Qy      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Db      3121 TGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCAT 3180
Qy      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
```

```
Db      3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCT 3240
Qy      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Db      3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCA 3300
Qy      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
Db      3301 TTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGG 3360
Qy      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
Db      3361 ATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGATGTCAGTGCATCAG 3420
Qy      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCCATATTCAGTGT 3480
Db      3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTTTGTCCCCCATATTCAGTGT 3480
Qy      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
Db      3481 CTGTAGATAAGAGACATGAAATGTCCCAAGTTTCTTTTGATCCATGGTGAACCTACTTAA 3540
Qy      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
Db      3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGCGGCAATCCTTGTGCAGTTGGGCCG 3600
Qy      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
Db      3601 TGGAATACAGCATATGCAAGTGTTTGATTGTGCAGCATTCTTTATTACTTGGTCGCAATA 3660
Qy      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
Db      3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTTGATTCTGCACTGCTCCCGTGTACAAACT 3720
Qy      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
Db      3721 TGGTTCTCAATAAGGCAGGCAGGAATGCATCTGCCAGTGGAACAGAGCAACCTGCACATT 3780
Qy      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
Db      3781 ATTTATGTATGCCTGTTCATTGGAGGGCTTGTTCATTACATGTTCGTCTATACTAGAAAA 3840
Qy      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
Db      3841 AACAGAATATTAGCATTAATCTATAGTTAATTAAAGTATGTAAATGCGCCTGTTTTTTGT 3900
Qy      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
Db      3901 TGTGTACTGTAATCATCTGAGTTGGTTTTGTGAAAA 3936
```

RESULT 5
US-09-900-237-3
; Sequence 3, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 3776
; TYPE: DNA
; ORGANISM: Zea mays

US-09-900-237-3

Query Match 65.9%; Score 2592.6; DB 10; Length 3776;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2885; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

```
Qy       104 CTCTGCCTTCGTCGGCGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCT 163
Db       139 CGCTTGCCCCGGGATCTGGAGATCTGGTAGCGCCAGGGGGATGGAGGCCAGCGCCGGGCT 198

Qy       164 GGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGG 223
Db       199 GGTCGCCGGCTCGCACAACCGGAACGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGG 258

Qy       224 GCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGG 283
Db       259 GCCGAAGCCCATGGACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGG 318

Qy       284 CCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCG 343
Db       319 GCGCAACCCCGACGGGGAGCCGTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCG 378

Qy       344 GGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCG 403
Db       379 GGACTGCTACGAGTACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCG 438

Qy       404 ATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGA 463
Db       439 CTTCAAGCGCCTCAAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGA 498

Qy       464 TGACCTGGACAACGAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTC 520
Db       499 CGACCTGGAGAACGAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTC 558

Qy       521 CATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGC 580
Db       559 CATGCTCCACGCCCACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCC 618

Qy       581 TTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCC 640
Db       619 ATTCCACCCCATCCCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCC 678

Qy       641 ACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGGAAAGAGGATACATCC 700
Db       679 GCCGGACCAGCACGCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGAAGAGGATTCACCC 738

Qy       701 CCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGA 760
Db       739 TCTCCCGTACGCGGATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGA 798

Qy       761 TCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCA 820
Db       799 TCTCGCCGCATATGGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCA 858

Qy       821 GAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGC 880
Db       859 GAAGCAGGAGAGGATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGC 918

Qy       881 TGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATC 940
Db       919 AGATCTACCACTAATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTC 978

Qy       941 AAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTT 1000
Db       979 AAGCCAAATCAACCCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTT 1038

Qy      1001 CTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGT 1060
Db      1039 CTTCCACTACCGAGTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGT 1098

Qy      1061 TATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCC 1120
Db      1099 GATCTGTGAAATTTGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCC 1158

Qy      1121 TATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCC 1180
Db      1159 TATCGAGAGGGAAACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCC 1218

Qy      1181 ATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCC 1240
Db      1219 TTCTCAACTCGCCCCTGTTGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCC 1278

Qy      1241 TTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGT 1300
Db      1279 ATTGGTCACTGCTAATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGT 1338

Qy      1301 TTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAAC 1360
Db      1339 TTCATGCTACGTTTCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAAC 1398

Qy      1361 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGC 1420
Db      1399 ATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGC 1458

Qy      1421 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTT 1480
Db      1459 TCCAGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTT 1518

Qy      1481 TGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGC 1540
Db      1519 TGTTAGAGAACGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGC 1578

Qy      1541 CTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCC 1600
Db      1579 CTTGGTTGCTAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCC 1638

Qy      1601 CTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAG 1660
Db      1639 ATGGCCCGGAAATAATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAG 1698

Qy      1661 CGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAA 1720
Db      1699 TGGTGGCCATGATGTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAA 1758

Qy      1721 ACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1780
Db      1759 ACGGCCAGGCTACAACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTC 1818

Qy      1781 TGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAA 1840
Db      1819 TGCTGTACTAACTAATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAA 1878

Qy      1841 CAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGT 1900
Db      1879 TAGTAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAAGAAAGT 1938

Qy      1901 TTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAA 1960
Db      1939 TTGCTATGTGCAGTTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAA 1998

Qy      1961 CCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAAT 2020
Db      1999 CAGAAATGTTGTCTTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAAT 2058

Qy      2021 TTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAA 2080
Db      2059 TTATGTGGGTACTGGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAA 2118

Qy      2081 AACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTG 2140
Db      2119 AACAAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTG 2178

Qy      2141 CTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAA 2194
Db      2179 CTGTTTTGGTAACAGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAA 2238
```

```
Qy       2195 AAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAAT 2254
Db       2239 GATAAAGAAACTTTTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAAT 2298

Qy       2255 TGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATT 2314
Db       2299 TGATGAAGCCGCTCCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTT 2358

Qy       2315 AGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGG 2374
Db       2359 GGAAAAGAAATTTGGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGG 2418

Qy       2375 AACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTG 2434
Db       2419 AACCCTGAAGAGTGCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGTTG 2478

Qy       2435 TGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTAC 2494
Db       2479 TGGATATGAAGACAAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCAC 2538

Qy       2495 AGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCAT 2554
Db       2539 AGAAGATATTCTTACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCAT 2598

Qy       2555 ACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCA 2614
Db       2599 ACCTAAACGGGCCGCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTCTTCACCA 2658

Qy       2615 GGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTG 2674
Db       2659 GGTTCTTCGGTGGGCTCTTGGTTCAATTGAAATTTTCTTCAGCAACCACTGCCCTCTCTG 2718

Qy       2675 GTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGT 2734
Db       2719 GTATGGGTATGGTGGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGT 2778

Qy       2735 GTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATT 2794
Db       2779 ATACCCTTGGACATCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCT 2838

Qy       2795 GACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACT 2854
Db       2839 GACAGGGAAATTTATCACGCCAGAGCTTAACAATGTTGCCAGCCTCTGGTTCATGTCACT 2898

Qy       2855 TTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGA 2914
Db       2899 TTTCATCTGCATTTTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGA 2958

Qy       2915 TGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGC 2974
Db       2959 TGACTGGTGGAGAAACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGC 3018

Qy       2975 TGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATC 3034
Db       3019 TGTGTTCCAGGGACTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATC 3078

Qy       3035 AAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATT 3094
Db       3079 CAAGGGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCT 3138

Qy       3095 GATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAA 3154
Db       3139 GATACCTCCGACAACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAA 3198

Qy       3155 TGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATT 3214
Db       3199 TGCGATCAACAACGGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATT 3258

Qy       3215 TTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGAC 3274
Db       3259 TTGGGTGATCGTCCATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGAC 3318

Qy       3275 ACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGT 3334
Db       3319 GCCAACGATTGTCATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGT 3378

Qy       3335 TCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGA 3394
Db       3379 CCGGATCGACCCGTTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGA 3438

Qy       3395 TTGCAACTAGGATGTCAGTGCATCAGCTCCCCC 3427
Db       3439 TTGCAACTAGGAGGTCAGCACGTGGACTTCCCC 3471
```

RESULT 6
US-10-209-059-41
; Sequence 41, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3487
; OTHER INFORMATION: n = A,T,C or G
US-10-209-059-41

Query Match 64.7%; Score 2546.8; DB 13; Length 3568;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

```
Qy        177 CACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCG 236
Db          6 CCCACGCGTCCGGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGGGCCGAAGCCCATG 65

Qy        237 CGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
Db         66 GACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGGGCGCAACCCCGAC 125

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
Db        126 GGGGAGCCTTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCGGGACTGCTACGAG 185

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
Db        186 TACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCGCTTCAAGCGCTTC 245

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
Db        246 AAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGACGACCTGGAGAAC 305

Qy        477 GAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGC 533
Db        306 GAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTCCATGCTCCACGCC 365

Qy        534 CACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAAC 593
Db        366 CACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCCATTCCACCCCATC 425
```

```
Qy       594 CCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCAC 653
Db       426 CCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCCGCCGGACCAGCAC 485

Qy       654 GCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGGAAAGAGGATACATCCCCTTCCTTATGCG 713
Db       486 GCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGAAGAGGATTCACCCTCTCCCGTACGCG 545

Qy       714 GATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATAT 773
Db       546 GATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGATCTCGCCGCATAT 605

Qy       774 GGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGG 833
Db       606 GGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCAGAAGCAGGAGAGG 665

Qy       834 ATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTA 893
Db       666 ATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGCAGATCTACCACTA 725

Qy       894 ATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAAT 953
Db       726 ATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATCAAC 785

Qy       954 CCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGA 1013
Db       786 CCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTTCTTCCACTACCGA 845

Qy      1014 GTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATC 1073
Db       846 GTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGTGATCTGTGAAATT 905

Qy      1074 TGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAG 1133
Db       906 TGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCCTATCGAGAGGGAA 965

Qy      1134 ACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCT 1193
Db       966 ACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCCTTCTCAACTCGCC 1025

Qy      1194 CCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACA 1253
Db      1026 CCTGTTGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCCATTGGTCACTGCT 1085

Qy      1254 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTT 1313
Db      1086 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGTTTCATGCTACGTT 1145

Qy      1314 TCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCA 1373
Db      1146 TCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAACATCTGAATTTGCA 1205

Qy      1374 AAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTAC 1433
Db      1206 AAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGCTCCAGAGTGGTAC 1265

Qy      1434 TTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGG 1493
Db      1266 TTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTTTGTTAGAGAACGG 1325

Qy      1494 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAA 1553
Db      1326 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGCCTTGGTTGCTAAA 1385

Qy      1554 GCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAAC 1613
Db      1386 GCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCCATGGCCCGGAAAT 1445

Qy      1614 AATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGAC 1673
Db      1446 AATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAGTGGTGGCCATGAT 1505
```

```
Qy      1674 TGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTAT 1733
Db      1506 GTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAAACGGCCAGGCTAC 1565

Qy      1734 AACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACA 1793
Db      1566 AACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACT 1625

Qy      1794 AATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATA 1853
Db      1626 AATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAATAGTAAGGCTATA 1685

Qy      1854 AAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAG 1913
Db      1686 AAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAAGAAAGTTTGCTATGTGCAG 1745

Qy      1914 TTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTC 1973
Db      1746 TTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAACAGAAATGTTGTC 1805

Qy      1974 TTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACT 2033
Db      1806 TTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAATTTATGTGGGTACT 1865

Qy      2034 GGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCA 2093
Db      1866 GGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAAAACAAAGAAGCCA 1925

Qy      2094 CCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAAT 2153
Db      1926 CCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTTGGTAAC 1985

Qy      2154 AGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAAAAAGTTATTATTT 2207
Db      1986 AGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAAGATAAAGAAACTT 2045

Qy      2208 TTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCT 2267
Db      2046 TTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAATTGATGAAGCCGCT 2105

Qy      2268 CCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTT 2327
Db      2106 CCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTTGGAAAAGAAATTT 2165

Qy      2328 GGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGT 2387
Db      2166 GGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGGAACCCTGAAGAGT 2225

Qy      2388 GCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGAC 2447
Db      2226 GCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGTTGTGGATATGAAGAC 2285

Qy      2448 AAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTA 2507
Db      2286 AAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCACAGAAGATATTCTT 2345

Qy      2508 ACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTT 2567
Db      2346 ACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGCC 2405

Qy      2568 GCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGG 2627
Db      2406 GCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTTTTCACCAGGTTCTTCGGTGG 2465

Qy      2628 GCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGT 2687
Db      2466 GCTCTTGGTTCAATTGAAATTTTGTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGT 2525

Qy      2688 GGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACA 2747
Db      2526 GGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGTATACCCTTGGACA 2585

Qy      2748 TCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTT 2807
```

```
Db       2586 TCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCTGACAGGGAAATTT 2645

Qy       2808 ATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATT 2867

Db       2646 ATCACGCCAGAGCTTAACAATGTTGCCAGCCTCTGGTTCATGTCACTTTTCATCTGCATT 2705

Qy       2868 TTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGG 2927

Db       2706 TTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGATGACTGGTGGAGA 2765

Qy       2928 AATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGA 2987

Db       2766 AACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGCTGTGTTCCAGGGA 2825

Qy       2988 CTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGAT 3047

Db       2826 CTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAGGGCGGAGAC 2885

Qy       3048 GATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACC 3107

Db       2886 GACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCTGATACCTCCGACA 2945

Qy       3108 ACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAAC 3167

Db       2946 ACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAAC 3005

Qy       3168 GGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTC 3227

Db       3006 GGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATTTTGGGTGATCGTC 3065

Qy       3228 CATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTC 3287

Db       3066 CATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTGTC 3125

Qy       3288 ATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCT 3347

Db       3126 ATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGTCCGGATCGACCCG 3185

Qy       3348 TTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGAT 3407

Db       3186 TTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGATTGCAACTAGGAG 3245

Qy       3408 GTCAGTGCATCAGCTCCCCC 3427

Db       3246 GTCAGCACGTGGACTTCCCC 3265


RESULT 7
US-10-160-719-1
; Sequence 1, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
;  APPLICANT: Kanwarpal S. Dhugga
;  APPLICANT:  Timothy G. Helentjaris
;  APPLICANT:  Benjamin A.  Bowen
;  APPLICANT:  Xun Wang
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864
;  CURRENT APPLICATION NUMBER: US/10/160,719
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/371,383A
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: US 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 3568
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (63)...(3239)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3487)...(3487)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(3568)
;   OTHER INFORMATION: n = A,T,C or G
US-10-160-719-1

  Query Match             64.7%;  Score 2546.8;  DB 13;  Length 3568;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches 2829;  Conservative    0;  Mismatches  422;  Indels    9;  Gaps    2;

Qy        177 CACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCG 236

Db          6 CCCACGCGTCCGGAGCTCGTCGTCATCCGCCGCGATGGCGAGCCAGGGCCGAAGCCCATG 65

Qy        237 CGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296

Db         66 GACCAGCGGAACGGCCAGGTGTGCCAGATTTGCGGCGACGACGTGGGGCGCAACCCCGAC 125

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356

Db        126 GGGGAGCCTTTCGTGGCCTGCAACGAGTGCGCCTTCCCCATCTGCCGGGACTGCTACGAG 185

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416

Db        186 TACGAGCGCCGCGAGGGCACGCAGAACTGCCCCCAGTGCAAGACCCGCTTCAAGCGCTTC 245

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476

Db        246 AAGGGGTGCGCGCGCGTGCCCGGGGACGAGGAGGAGGACGGCGTCGACGACCTGGAGAAC 305

Qy        477 GAGTTCAACTGG---GACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGC 533

Db        306 GAGTTCAACTGGAGCGACAAGCACGACTCCCAGTACCTCGCCGAGTCCATGCTCCACGCC 365

Qy        534 CACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAAC 593

Db        366 CACATGAGCTACGGCCGCGGCGCCGACCTCGACGGCGTGCCGCAGCCATTCCACCCCATC 425

Qy        594 CCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCAC 653

Db        426 CCCAATGTTCCCCTCCTCACCAACGGACAGATGGTCGATGACATCCCGCCGGACCAGCAC 485

Qy        654 GCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCG 713

Db        486 GCCCTTGTGCCCTCGTTCGTGGGTGGCGGGGGGAAGAGGATTCACCCTCTCCCGTACGCG 545

Qy        714 GATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATAT 773

Db        546 GATCCCAACCTTCCTGTGCAACCGAGGTCTATGGACCCTTCCAAGGATCTCGCCGCATAT 605

Qy        774 GGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGG 833

Db        606 GGCTACGGGAGCGTAGCATGGAAGGAGAGGATGGAGAGCTGGAAGCAGAAGCAGGAGAGG 665

Qy        834 ATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTA 893

Db        666 ATGCACCAGACGAGGAACGATGGCGGCGGCGATGATGGTGATGATGCAGATCTACCACTA 725

Qy        894 ATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAAT 953

Db        726 ATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATCAAC 785

Qy        954 CCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGA 1013

Db        786 CCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTTGTGTTTCTTCTTCCACTACCGA 845

Qy       1014 GTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATC 1073
```

```
Db        846 GTGATGCATCCGGTGCCTGATGCATTTGCTTTATGGCTCATATCTGTGATCTGTGAAATT 905

Qy       1074 TGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAG 1133

Db        906 TGGTTTGCCATGTCTTGGATTCTTGACCAGTTTCCAAAGTGGTTTCCTATCGAGAGGGAA 965

Qy       1134 ACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCT 1193

Db        966 ACCTATCTTGACCGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCCTTCTCAACTCGCC 1025

Qy       1194 CCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACA 1253

Db       1026 CCTGTTGATTTCTTTGTCAGTACGGTTGATCCCTTGAAGGAACCTCCATTGGTCACTGCT 1085

Qy       1254 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTT 1313

Db       1086 AATACTGTTCTATCTATCCTTTCGGTGGATTATCCAGTTGATAAGGTTTCATGCTACGTT 1145

Qy       1314 TCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCA 1373

Db       1146 TCTGATGATGGTGCTGCCATGCTGACATTTGAAGCATTGTCTGAAACATCTGAATTTGCA 1205

Qy       1374 AAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTAC 1433

Db       1206 AAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCTTGAGCCTCGTGCTCCAGAGTGGTAC 1265

Qy       1434 TTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGG 1493

Db       1266 TTCCAACAGAAGATAGACTACCTGAAAGACAAGGTGGCGCCAAACTTTGTTAGAGAACGG 1325

Qy       1494 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAA 1553

Db       1326 AGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTCAGAATCAATGCCTTGGTTGCTAAA 1385

Qy       1554 GCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAAC 1613

Db       1386 GCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAGGATGGAACTCCATGGCCCGGAAAT 1445

Qy       1614 AATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGAC 1673

Db       1446 AATGTCCGTGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAAAGTGGTGGCCATGAT 1505

Qy       1674 TGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTAT 1733

Db       1506 GTGGAAGGAAATGAGCTGCCTCGATTGGTTTATGTTTCAAGAGAAAAACGGCCAGGCTAC 1565

Qy       1734 AACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACA 1793

Db       1566 AACCATCACAAGAAGGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACT 1625

Qy       1794 AATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATA 1853

Db       1626 AATGCTCCTTATTTGCTGAACTTGGATTGTGATCACTATATCAATAATAGTAAGGCTATA 1685

Qy       1854 AAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAG 1913

Db       1686 AAGGAAGCAATGTGTTTTATGATGGATCCTTTGCTTGGAAAGAAAGTTTGCTATGTGCAG 1745

Qy       1914 TTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTC 1973

Db       1746 TTTCCTCAAAGATTTGATGGGATTGATCGCCATGATCGATATGCTAACAGAAATGTTGTC 1805

Qy       1974 TTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACT 2033

Db       1806 TTTTTCGATATCAACATGAAAGGTTTGGATGGTATCCAGGGCCCAATTTATGTGGGTACT 1865

Qy       2034 GGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCA 2093

Db       1866 GGATGTGTCTTCAGAAGGCAGGCATTATATGGCTACGATGCTCCCAAAACAAAGAAGCCA 1925

Qy       2094 CCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAAT 2153

Db       1926 CCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTTGGTAAC 1985
```

```
Qy       2154 AGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGA------AGAAAAAGTTATTATTT 2207

Db       1986 AGGAAGACCAAGAAGAAGACCAAGACCTCTAAACCTAAATTTGAGAAGATAAAGAAACTT 2045

Qy       2208 TTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCT 2267

Db       2046 TTTAAGAAAAAGGAAAATCAAGCCCCTGCATATGCTCTTGGTGAAATTGATGAAGCCGCT 2105

Qy       2268 CCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTT 2327

Db       2106 CCAGGAGCTGAAAATGAAAAGGCTAGTATTGTAAATCAACAGAAGTTGGAAAAGAAATTT 2165

Qy       2328 GGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGT 2387

Db       2166 GGCCAGTCTTCAGTTTTTGTTGCATCCACACTTCTTGAGAATGGTGGAACCCTGAAGAGT 2225

Qy       2388 GCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGAC 2447

Db       2226 GCCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTCATCAGTTGTGGATATGAAGAC 2285

Qy       2448 AAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTA 2507

Db       2286 AAAACAGGCTGGGGAAAAGATATTGGTTGGATTTATGGATCAGTCACAGAAGATATTCTT 2345

Qy       2508 ACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTT 2567

Db       2346 ACTGGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGCC 2405

Qy       2568 GCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGG 2627

Db       2406 GCCTTCAAAGGTTCCGCACCTCTCAATCTTTCCGATCGTTTTCACCAGGTTCTTCGGTGG 2465

Qy       2628 GCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGT 2687

Db       2466 GCTCTTGGTTCAATTGAAATTTTGTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGT 2525

Qy       2688 GGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACA 2747

Db       2526 GGTGGACTAAAGTTCCTGGAAAGGTTTTCGTACATTAACTCCATCGTATACCCTTGGACA 2585

Qy       2748 TCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTT 2807

Db       2586 TCTATCCCGCTCTTGGCCTATTGCACATTGCCTGCCATCTGCTTGCTGACAGGGAAATTT 2645

Qy       2808 ATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATT 2867

Db       2646 ATCACGCCAGAGCTTAACAATGTTGCCAGCCTCTGGTTCATGTCACTTTTCATCTGCATT 2705

Qy       2868 TTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGG 2927

Db       2706 TTTGCTACGAGCATCCTGGAAATGAGATGGAGTGGTGTAGGCATCGATGACTGGTGGAGA 2765

Qy       2928 AATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGA 2987

Db       2766 AACGAGCAGTTTTGGGTCATTGGAGGCGTGTCTTCACATCTCTTTGCTGTGTTCCAGGGA 2825

Qy       2988 CTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGAT 3047

Db       2826 CTCCTCAAGGTCATAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAGGGCGGAGAC 2885

Qy       3048 GATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACC 3107

Db       2886 GACGAGGAGTTCTCAGAGCTGTACACATTCAAATGGACGACCCTTCTGATACCTCCGACA 2945

Qy       3108 ACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAAC 3167

Db       2946 ACCCTGCTCCTACTGAACTTCATTGGAGTGGTAGCTGGCATCTCCAATGCGATCAACAAC 3005

Qy       3168 GGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTC 3227

Db       3006 GGATATGAATCATGGGGCCCCCTGTTCGGGAAGCTCTTCTTTGCATTTTGGGTGATCGTC 3065
```

```
Qy       3228 CATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTC 3287
             |||||||| || || ||||| ||| ||||||| ||||| |||||||| ||||||||||||
Db       3066 CATCTTTACCCGTTCCTCAAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTGTC 3125

Qy       3288 ATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCT 3347
             || |||||||||||| || |||||||| |||||||||| |||||||| ||||| || ||
Db       3126 ATTGTCTGGTCCATCCTCCTGGCTTCGATCTTCTCGCTGCTTTGGGTCCGGATCGACCCG 3185

Qy       3348 TTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGCAACTAGGAT 3407
             |||||||||||||||||||||||| ||   | |||||||||||| ||||||||||||||||
Db       3186 TTCCTTGCGAAGGATGATGGTCCCCTGTTGGAGGAGTGTGGTCTGGATTGCAACTAGGAG 3245

Qy       3408 GTCAGTGCATCAGCTCCCCC 3427
             |||||  | |   || ||||
Db       3246 GTCAGCACGTGGACTTCCCC 3265
```

RESULT 8
US-09-900-237-7
; Sequence 7, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 3786
; TYPE: DNA
; ORGANISM: Zea mays
US-09-900-237-7

Query Match 56.9%; Score 2241; DB 10; Length 3786;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| |||||||||||||||||| || |||
Db        194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || ||||||||| || || |||||||||||| |||| ||  | ||||  | || ||  |
Db        254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||    |||||       | ||||||||| |||||||||| |||||  | | |  || ||
Db        314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||||||| | |||||||||
Db        374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||||  |||||| ||||     ||||| ||||||| ||| || ||||||||||||
Db        434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              ||||||||||   || ||| |||| |||||||||||||||||||||| ||||||||    |
Db        494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy        477 GAGTTCAACTGG--------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              ||||||  |  |              |||   | ||| ||||||  || ||||||||||
Db        554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
              ||||||  |  || |||||||||||||| || || ||          |||||  |  |
Db        614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
              |||  | |  |||||| || || |||||||||||||| || ||||| ||||||||||||
Db        665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
              |||||||||||||||||| ||||| || | ||||         || || ||||| ||||||
Db        725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
              || || || || |||| || ||||||| | | || |||||||| || || ||||||||
Db        785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
              |||||||||||  || || || || ||||| || || || ||||||||  | ||||||
Db        845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
              |||||||||  || ||| |  |||| ||     |||  || ||||| ||||||   |||
Db        905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
              || |||||||| |||||  |||||||||||||||||| || || | | ||||||| ||| |
Db        965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
              ||  |  |||||||||  |||||||| || |||||||||||| |||| ||| | ||||||
Db       1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084

Qy        990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
              ||||| |||||||||||||||||||||||||||||||||| ||| |||||||||| ||||||
Db       1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy       1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109
              |||||||||||  |||||||||||||||||| ||||| |||||||||||||||| ||||||
Db       1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy       1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
              ||||||  | ||  || |||||||||||||||| |||||  ||||||| |||||  ||||||
Db       1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy       1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
              |||||  || ||  |||||  |||||||||| ||  |||||||||||||||||||||||  |
Db       1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy       1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
              ||||||||||  |||||||||  | ||  |||||  || ||  |||||||| ||||||||||
Db       1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy       1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
              |||||  |||||  ||||||||||||||||||||||||||||||||||  ||||||||||
Db       1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy       1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
              || |||||||||||||||||||||||||||||||||||||||||  |||||  ||  |||||
Db       1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy       1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
              || |||||  |||||  |||||||||||||||||||||||||||||||  ||||||||||||
Db       1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy       1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
              || ||    |||||||||||||||| || ||||||||||  || ||||||||||||||
Db       1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy       1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
              || |||||||||| |||||  |||||||| || |||||||| ||||||||||||||||||
Db       1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy       1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
```

```
Db       1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy       1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709

Db       1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy       1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769

Db       1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy       1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829

Db       1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy       1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889

Db       1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy       1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949

Db       1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy       1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009

Db       2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy       2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069

Db       2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy       2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129

Db       2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy       2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189

Db       2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy       2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249

Db       2285 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344

Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309

Db       2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369

Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429

Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489

Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549

Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609

Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669

Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729

Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789

Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849

Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909

Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969

Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029

Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089

Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149

Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209

Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269

Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329

Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389

Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402

Db       3485 CTGGACTGCAACT 3497
```

RESULT 9
US-10-209-059-17
; Sequence 17, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3813

; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-17

Query Match 56.9%; Score 2241; DB 13; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
Db        194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
Db        254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
Db        314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
Db        374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
Db        434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
Db        494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy        477 GAGTTCAACTGG----------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
Db        554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
Db        614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
Db        665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
Db        725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
Db        785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
Db        845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
Db        905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
Db        965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
Db       1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084

Qy        990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
Db       1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy       1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109

Db       1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy       1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
Db       1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy       1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
Db       1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy       1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
Db       1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy       1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
Db       1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy       1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
Db       1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy       1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
Db       1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy       1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
Db       1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy       1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
Db       1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy       1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
Db       1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy       1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709
Db       1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy       1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769
Db       1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy       1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829
Db       1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy       1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889
Db       1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy       1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949
Db       1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy       1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009
Db       2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy       2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069
Db       2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy       2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129
Db       2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy       2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189
```

```
Db       2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy       2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249
             ||||| ||   |||||||||||||||||| ||| |||| ||| |||||||||||| | |||
Db       2285 AAGAAGAAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344

Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309
             |||||||| |||| ||||||||| ||||| |  || |||||||| || ||||||||||
Db       2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369
             ||| |||| |||||||||| || |||||||||||||  |||| |||||||| |||||
Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429
             || || ||| ||||||| |||||||| ||||||||| |||| |||||||||||||| ||
Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489
             || || || || |||||||||| ||||||||||||||||||||||||| || |||||
Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549
              | ||||| |||||  | ||||| || |||||||| || ||||| |||||||| ||||||
Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609
             ||||| || || |||  |||||||||||||||||| |||||||| ||||| || ||||||
Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669
             |||||||| ||||| ||||| |||||||||  | || || |||||||||| || |||||
Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729
             |||||||| || || || ||||| || |||||  ||||||| ||||| || |||||||||
Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
             ||||| ||||| ||||| || ||||| ||| |||||||||||| ||||||||||||||
Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
              |  | || |||||| |||||||| |||||||| |  ||||| |||||  | |||||||||
Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
              ||||||| |||||||| |   |  ||  ||||||| |||||||| |||||||| || |
Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
             || || |||||||||||||| |||||||||||||||| ||||| || ||  | || ||
Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
             || || |||||||||||  || || |||||  | || ||  | || || ||||||||||||
Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
             || || ||||  || || || |||||||||||| ||||| || || |||||  ||||| |||
Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
              |  ||||||| || |||||  | || ||  |||||||||| || ||||| || ||   |
Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
             || || |||||||| ||||| || ||||| |||||||||| || ||||||||  |||||
Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
             || || |||||||| ||||| || || || || ||||| ||| |||| || |||||  |||
Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364
```

```
Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
             ||||| || ||||| |||||||||||||||| |||||||| || ||||||||||||||
Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
             ||||| ||  | || || ||||| || |||    | || ||||| || ||||||||||
Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402
              |||| |||||||
Db       3485 CTGGACTGCAACT 3497
```

RESULT 10
US-10-160-719-25
; Sequence 25, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 3813
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(3496)
US-10-160-719-25

Query Match 56.9%; Score 2241; DB 13; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| |||||||||||||||| || |||
Db        194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || |||||||||| || || |||||||||| ||  | ||||  | || ||  |
Db        254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||    |||||       | ||||||||| |||||||||| |||||  | | |  || |
Db        314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||||||| | |||||||||
Db        374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||||  |||||| ||||     |||||  ||||||| ||| || ||||||||||||
Db        434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              ||||||||||   || ||| |||| |||||||||||||||||||||||| |||||||   |
Db        494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy        477 GAGTTCAACTGG---------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              ||||||  |  |                |||   | ||| |||||||  || ||||||||||
```

```
Db        554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581

Db        614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641

Db        665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692

Db        725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752

Db        785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812

Db        845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869

Db        905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929

Db        965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989

Db       1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084

Qy        990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049

Db       1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy       1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109

Db       1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy       1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169

Db       1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy       1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229

Db       1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy       1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289

Db       1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy       1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349

Db       1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy       1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409

Db       1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy       1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469

Db       1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy       1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529

Db       1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy       1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589

Db       1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy       1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649

Db       1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy       1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709

Db       1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy       1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769

Db       1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy       1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829

Db       1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy       1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889

Db       1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy       1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949

Db       1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy       1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009

Db       2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy       2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069

Db       2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy       2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129

Db       2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy       2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189

Db       2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy       2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249

Db       2285 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344

Qy       2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309

Db       2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy       2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369

Db       2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy       2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429

Db       2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy       2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489

Db       2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy       2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549

Db       2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy       2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609

Db       2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy       2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669

Db       2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764
```

```
Qy       2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729
             |||||||| || || || ||||| || |||||  ||||||| ||||| || |||||||||
Db       2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
             ||||| ||||| ||||| || ||||| ||| |||||||||||| ||||||||||||||
Db       2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
               | | || ||||| |||||||| |||||||| | ||||| |||||  | |||||||||
Db       2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
              ||||||| |||||||| | |  ||  ||||||| |||||||| |||||||| || |
Db       2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
              || || |||||||||||||| |||||||||||||||| ||||| || ||  | || ||
Db       3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
              || || |||||||||||  || || |||||  | || ||  | || || ||||||||||||
Db       3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
              || ||  ||||  || || || |||||||||| ||||| || || ||||| |||||  |||
Db       3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
               |  ||||||| || |||||  | || ||  |||||||||| || ||||| || ||  |
Db       3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
              || || |||||||| ||||| || ||||| |||||||||| || |||||||| |||||
Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
              || || |||||||| ||||| || || || || ||||| ||| |||| || ||||| |||
Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
              ||||| || ||||| |||||||||||||||| |||||||| || ||||||||||||||
Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389
              ||||| ||  | || || ||||| || |||    | || ||||| || ||||||||||
Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402
              |||| |||||||
Db       3485 CTGGACTGCAACT 3497
```

RESULT 11
US-10-160-719-45
; Sequence 45, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 45
; LENGTH: 3813
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(3496)
US-10-160-719-45

Query Match 56.9%; Score 2241; DB 13; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

```
Qy        123 GGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAAC 182
              ||  | || |||| |   | ||||||||| ||||| |||||||||||||||||| || |||
Db        194 GGGCGGGGCGGGGGGCCTGAGATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAC 253

Qy        183 CGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA-TCCCGGGCCGAAGCCGCCGCGGGA 241
              || |||||||| || || |||||||||||| |||| ||  | ||||  | || ||  |
Db        254 CGGAACGAGCTGGTGGTGATCCGCCGCGACCGCGAGTCGGGAGCCGCGGGCGGCGGCGCG 313

Qy        242 GCAGAACGGGC-----AGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGC 296
              ||    |||||       | ||||||||| |||||||||| |||||  | | |  || |
Db        314 GCGCGCCGGGCGGAGGCGCCGTGCCAGATATGCGGCGACGAGGTCGGGGTGGGCTTCGAC 373

Qy        297 GGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAA 356
              ||||| |||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db        374 GGGGAGCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGCGCCTGCTACGAG 433

Qy        357 TACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTC 416
              |||||||||||  || ||| |||||     ||||| |||||||  ||| || ||||||||||
Db        434 TACGAGCGCCGCGAGGGCTCGCAAGCGTGCCCGCAGTGCAGGACCCGCTACAAGCGCCTC 493

Qy        417 AAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAAC 476
              ||||||||||  || ||| |||| |||||||||||||||||||||||||| |||||||  |
Db        494 AAGGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTCGACGACCTGGAGGGC 553

Qy        477 GAGTTCAACTGG---------------GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521
              ||||||  |  |                |||   | ||| ||||||  || |||||||||
Db        554 GAGTTCGGCCTGCAGGACGGCGCCGCCCACGAGGACGACCCGCAGTACGTCGCCGAGTCC 613

Qy        522 ATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCT 581
              ||||||   |  || ||||||||||||| || || ||         |||||  |  |
Db        614 ATGCTCAGGGCGCAGATGAGCTACGGCCGCGGCGGCGA---------CGCGCACCCCGGC 664

Qy        582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCA 641
              |||   | |  |||||| || || |||||||||||||| || ||||| ||||||||||
Db        665 TTCAGCCCCGTCCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724

Qy        642 CCGGAGCAGCACGCGCTGGTGCCTTCTTTCATG---------GGTGGTGGGGGAAAGAGG 692
              |||||||||||||||| ||||| || | ||||         || || |||||  ||||||
Db        725 CCGGAGCAGCACGCGCTCGTGCCGTCCTACATGAGCGGCGGCGGCGGCGGGGGCAAGAGG 784

Qy        693 ATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCA 752
              || || || || ||||  || ||||||| | | || |||||||| || || ||||||||
Db        785 ATCCACCCGCTCCCTTTCGCAGATCCCAACCTTCCAGTGCAACCGAGATCCATGGACCCG 844

Qy        753 TCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAAT 812
              |||||||||| || || || || ||||| || || || ||||||||  | ||||||
Db        845 TCCAAGGATCTGGCCGCCTACGGATATGGCAGCGTGGCCTGGAAGGAGAGAATGGAGGGC 904

Qy        813 TGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGAT---GAT 869
              |||||||||  || ||| |  |||| ||     |||  || ||||| ||||||   |||
Db        905 TGGAAGCAGAAGCAGGAGCGCCTGCAGCATGTCAGGAGCGAGGGTGGCGGTGATTGGGAT 964

Qy        870 GGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATT 929
              || |||||||| ||||| |||||||||||||||| || || | | ||||||| ||| |
Db        965 GGCGACGATGCAGATCTGCCACTAATGGATGAAGCTAGGCAGCCATTGTCCAGAAAAGTC 1024

Qy        930 CCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTT 989
              ||  |  ||||||||  |||||||| || |||||||||||| |||| ||| | ||||||
Db       1025 CCTATATCATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTT 1084
```

```
Qy       990 TTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGG 1049
Db      1085 TTGGGTTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAGATGCATTTGCATTGTGG 1144

Qy      1050 CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCA 1109
Db      1145 CTCATATCTGTAATCTGTGAAATCTGGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA 1204

Qy      1110 AAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAG 1169
Db      1205 AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGACAAG 1264

Qy      1170 GAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA 1229
Db      1265 GAAGGTCAACCCTCTCAGCTTGCTCCAATCGACTTCTTTGTCAGTACGGTTGATCCCACA 1324

Qy      1230 AAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCT 1289
Db      1325 AAGGAACCTCCCTTGGTCACAGCGAACACTGTCCTTTCCATCCTTTCTGTGGATTATCCG 1384

Qy      1290 GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCA 1349
Db      1385 GTTGAGAAGGTCTCCTGCTATGTTTCTGATGATGGTGCTGCAATGCTTACGTTTGAAGCA 1444

Qy      1350 TTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATT 1409
Db      1445 TTGTCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCAGCAAAAAGTTTAATATC 1504

Qy      1410 GAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTG 1469
Db      1505 GAGCCTCGTGCTCCTGAGTGGTACTTCCAACAGAAGATAGACTACCTGAAAGACAAGGTT 1564

Qy      1470 GCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTG 1529
Db      1565 GCTGCTTCATTTGTTAGGGAGAGGAGGGCGATGAAGAGAGAATACGAGGAATTCAAGGTA 1624

Qy      1530 AGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAA 1589
Db      1625 AGGATCAATGCCTTGGTTGCAAAAGCCCAAAAGGTTCCTGAGGAAGGATGGACAATGCAA 1684

Qy      1590 GATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTC 1649
Db      1685 GATGGAAGCCCCTGGCCTGGAAACAACGTACGCGATCATCCTGGAATGATTCAGGTATTC 1744

Qy      1650 CTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTT 1709
Db      1745 CTTGGCCAAAGTGGCGGTCGTGATGTGGAAGGAAATGAGTTGCCTCGCCTGGTTTATGTC 1804

Qy      1710 TCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTG 1769
Db      1805 TCGAGAGAAAAGAGGCCAGGTTATAACCATCACAAGAAGGCTGGTGCCATGAATGCACTG 1864

Qy      1770 GTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCAC 1829
Db      1865 GTCCGTGTCTCTGCTGTCTTATCAAATGCTGCATACCTATTGAACTTGGACTGTGATCAC 1924

Qy      1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTA 1889
Db      1925 TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTG 1984

Qy      1890 GGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGAC 1949
Db      1985 GGGAAGAAAGTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGGTATTGACAAAAATGAT 2044

Qy      1950 CGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATT 2009
Db      2045 CGATACGCTAACAGGAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATT 2104

Qy      2010 CAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTAT 2069
Db      2105 CAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGGTTAT 2164

Qy      2070 GATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGC 2129

Db      2165 GATGCTCCTAAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGC 2224

Qy      2130 TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAG 2189
Db      2225 CTCTCTTGCTGCTGCAGCAGGAACAAGAATAAAAAGAAGACTACAAAACCAAAGACGGAG 2284

Qy      2190 AAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGT 2249
Db      2285 AAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCTCCTGCATATGCTTTGGGT 2344

Qy      2250 GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA 2309
Db      2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGCCGGAATCGTAAATCAACAG 2404

Qy      2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAAT 2369
Db      2405 AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGCATCAACACTTCTTGAGAAC 2464

Qy      2370 GGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATT 2429
Db      2465 GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAAGGAAGCTATACATGTTATC 2524

Qy      2430 AGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCA 2489
Db      2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGGAAAAGAGATTGGCTGGATTTACGGATCG 2584

Qy      2490 GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTAC 2549
Db      2585 ATCACAGAGGATATCTTGACTGGATTTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy      2550 TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTT 2609
Db      2645 TGCATCCCGAAGCGGCCTGCATTCAAAGGTTCTGCGCCTCTGAACCTTTCCGACCGTCTT 2704

Qy      2610 CACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCT 2669
Db      2705 CACCAGGTCCTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCCA 2764

Qy      2670 CTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729
Db      2765 CTTTGGTACGGATACGGCGGCGGGCTAAAATTCCTGGAAAGGTTTTCTTATATCAACTCC 2824

Qy      2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
Db      2825 ATCGTTTATCCCTGGACGTCCATTCCTCTCCTGGCTTACTGTACCTTGCCTGCCATCTGC 2884

Qy      2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
Db      2885 CTGCTCACGGGGAAGTTTATCACACCAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG 2944

Qy      2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
Db      2945 GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGGCC 3004

Qy      2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
Db      3005 ATCGACGACTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCTG 3064

Qy      2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
Db      3065 TTCGCGGTGTTCCAGGGCCTGCTGAAGGTGTTCGCCGGCATCGACACGAGCTTCACCGTG 3124

Qy      3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
Db      3125 ACGTCGAAGGCCGGGGACGACGAGGAGTTCTCGGAGCTGTACACGTTCAAGTGGACCACC 3184

Qy      3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
Db      3185 CTGCTGATACCCCCGACCACGCTCCTCCTGCTGAACTTCATCGGGGTGGTGGCCGGGATC 3244

Qy      3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
```

```
Db       3245 TCGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269

Db       3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGGTCTGGTGGGGAGGCAGAAC 3364

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329

Db       3365 AGGACGCCGACGATCGTCATCGTCTGGTCCATCCTGCTGGCCTCGATCTTCTCGCTCCTG 3424

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGT 3389

Db       3425 TGGGTCCGCGTCGACCCGTTCCTCGCCAAGAGCAACGGCCCGCTCCTGGAGGAGTGTGGC 3484

Qy       3390 TTGGATTGCAACT 3402

Db       3485 CTGGACTGCAACT 3497
```

RESULT 12
US-09-938-842A-2530
; Sequence 2530, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2530
; LENGTH: 3255
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2530

Query Match 38.3%; Score 1509; DB 10; Length 3255;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 2207; Conservative 0; Mismatches 1030; Indels 30; Gaps 5;

```
Qy        144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203

Db          1 ATGAACACCGGTGGTCGGTTAATCGCCGGTTCTCACAACAGGAATGAGTTTGTCCTCATT 60

Qy        204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263

Db         61 AATGCCGATGAGAATGCCCGAATAAGATCAGTCCAAGAGCTGAGTGGACAGACATGTCAA 120

Qy        264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323

Db        121 ATCTGCAGAGATGAGATCGAATTGACTGTTGATGGAGAACCGTTTGTGGCATGTAACGAA 180

Qy        324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383

Db        181 TGTGCATTCCCTGTGTGTAGACCTTGCTATGAGTACGAAAGACGAGAAGGCAATCAAGCT 240

Qy        384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443

Db        241 TGTCCACAGTGCAAAACCCGTTTCAAACGTCTTAAAGGAAGTCCAAGAGTTGAAGGTGAT 300

Qy        444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503

Db        301 GAAGAGGAAGATGACATTGATGATTTAGACAATGAGTTTGA---GTATGGAAATAATGGG 357
```

```
Qy        504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563

Db        358 ATTGGATTTGATCAGGTTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCCA 417

Qy        564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623

Db        418 CAATCTGATTTGGATTCAGCTCCACCTGGCTCTCAGATTCCATTGCTGACTTACGGCGAC 477

Qy        624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGC---TGGTGCCTTCTTTCATGGGTGGT 680

Db        478 GAGGACGTTGAGATTTCTTCTGATAGACATGCTCTTATTGTTCCTCCTTCACTTGGTGGT 537

Qy        681 GGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGG 740

Db        538 CATGGCAATAGAGTTCATCCTGTTTCTCTTTCTGACCCGACCGTGGCTGCACATCCAAGG 597

Qy        741 TCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAA 800

Db        598 CCTATGGTACCTCAGAAAGATCTTGCGGTTTATGGTTATGGAAGTGTCGCTTGGAAAGAT 657

Qy        801 CGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGA---------AT 851

Db        658 CGGATGGAGGAATGGAAGAGAAAGCAGAATGAGAAACTTCAGGTTGTTAGGCATGAAGGA 717

Qy        852 GATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGAAGCAAGACAA 911

Db        718 GATCCTGATTTTGAAGATGGTGATGATGCTGATTTTCCAATGATGGATGAGGGAAGGCAG 777

Qy        912 CAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATC 971

Db        778 CCATTGTCTAGGAAGATACCAATCAAATCGAGCAAGATAAATCCTTACCGGATGTTAATT 837

Qy        972 ATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAAT 1031

Db        838 GTGCTACGTCTTGTGATTCTTGGTCTCTTCTTTCACTACCGTATTCTTCACCCCGTCAAA 897

Qy       1032 GATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGG 1091

Db        898 GATGCATATGCTTTGTGGCTTATTTCTGTTATATGTGAGATATGGTTTGCTGTTTCATGG 957

Qy       1092 ATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTG 1151

Db        958 GTTCTTGATCAGTTCCCTAAATGGTACCCTATCGAGCGAGAAACGTACTTGGACCGACTC 1017

Qy       1152 TCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTC 1211

Db       1018 TCATTAAGATATGAGAAAGAAGGGAAACCGTCGGGACTATCCCCTGTGGATGTATTTGTT 1077

Qy       1212 AGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATC 1271

Db       1078 AGTACAGTGGATCCATTGAAAGAGCCTCCGCTTATTACTGCAAATACTGTCTTGTCTATT 1137

Qy       1272 CTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCA 1331

Db       1138 CTTGCTGTTGATTATCCTGTCGATAAGGTTGCTTGTTACGTATCTGATGATGGTGCTGCT 1197

Qy       1332 ATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTC 1391

Db       1198 ATGCTTACTTTCGAAGCTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTC 1257

Qy       1392 TGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGAC 1451

Db       1258 TGCAAGAAATATTGTATTGAGCCTCGTGCTCCCGAATGGTATTTCTGCCATAAAATGGAC 1317

Qy       1452 TACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAG 1511

Db       1318 TACTTGAAGAATAAAGTTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGAT 1377

Qy       1512 TATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAA 1571

Db       1378 TATGAAGAATTCAAAGTAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCCTGAG 1437
```

```
Qy       1572 GAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCT 1631
Db       1438 GATGGTTGGACTATGCAAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCCT 1497

Qy       1632 GGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTG 1691
Db       1498 GGCATGATTCAGGTCTTCCTTGGAAGTGACGGTGTTCGTGATGTCGAAAACAACGAGTTG 1557

Qy       1692 CCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCT 1751
Db       1558 CCTCGATTAGTTTACGTTTCTCGTGAGAAGAGACCCGGATTTGATCACCATAAGAAGGCT 1617

Qy       1752 GGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTA 1811
Db       1618 GGAGCTATGAATTCCCTGATACGAGTCTCTGGGGTTCTATCAAATGCTCCTTACCTTCTG 1677

Qy       1812 AACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTT 1871
Db       1678 AATGTCGATTGTGATCACTACATCAACAATAGCAAAGCTCTTAGAGAAGCAATGTGTTTC 1737

Qy       1872 ATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGAT 1931
Db       1738 ATGATGGATCCTCAGTCAGGAAAGAAAATCTGTTATGTTCAGTTCCCTCAAAGGTTCGAT 1797

Qy       1932 GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATG 1991
Db       1798 GGGATTGATAGGCACGATCGATACTCAAATCGCAATGTTGTGTTCTTTGATATCAATATG 1857

Qy       1992 AAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGG 2051
Db       1858 AAAGGTTTGGATGGGCTACAAGGGCCTATATACGTCGGTACAGGTTGTGTTTTCAGGAGG 1917

Qy       2052 CAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAAC 2111
Db       1918 CAAGCGCTTTACGGATTTGATGCACCGAAGAAGAAGAAGGGCCCACGTAAGACATGCAAT 1977

Qy       2112 TGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACT 2171
Db       1978 TGCTGGCCAAAATGGTGTCTCCTATGTTTTGGTTCAAGAAAGAATCGTAAAGCAAAGACA 2037

Qy       2172 ACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCC 2231
Db       2038 GTGGCTGCGGATAAGAAGAAGAAGAA------------TAGGGAAGCGTCAAAGCAGATC 2085

Qy       2232 CCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCC 2291
Db       2086 CACGCATTAGAAAATATCGAAGAGGGCCGCGTCACTAAAGGTTCTAACGTAGAACAGTCA 2145

Qy       2292 GGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACA 2351
Db       2146 ACCGAGGCAATGCAAATGAAGTTGGAGAAGAAATTTGGGCAGTCTCCTGTATTTGTTGCA 2205

Qy       2352 TCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAA 2411
Db       2206 TCTGCGCGTATGGAGAATGGTGGGATGGCTAGAAACGCAAGCCCGGCTTGTCTGCTTAAA 2265

Qy       2412 GAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATT 2471
Db       2266 GAAGCCATCCAAGTCATTAGTTGCGGATATGAAGATAAAACTGAATGGGGAAAAGAGATT 2325

Qy       2472 GGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCAT 2531
Db       2326 GGGTGGATCTATGGTTCTGTTACCGAAGATATTCTTACGGGTTTTAAGATGCATTCTCAT 2385

Qy       2532 GGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTG 2591
Db       2386 GGTTGGAGATCTGTTTATTGTACACCAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATC 2445

Qy       2592 AATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651
Db       2446 AATCTTTCGGATCGTCTCCATCAAGTTCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTC 2505

Qy       2652 TTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGA 2711
```

```
Db       2506 TTGAGTAGGCATTGTCCTATTTGGTATGGTTATGGAGGTGGGTTGAAATGGCTTGAGCGG 2565

Qy       2712 TTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGT 2771
Db       2566 TTGTCCTACATTAACTCTGTGGTTTACCCGTGGACCTCTCTACCGCTCATCGTTTACTGT 2625

Qy       2772 ACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTT 2831
Db       2626 TCTCTCCCTGCCATCTGTCTTCTCACTGGAAAATTCATCGTTCCCGAGATTAGCAACTAT 2685

Qy       2832 GCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATG 2891
Db       2686 GCGAGTATCCTCTTCATGGCGCTCTTCTCGTCGATTGCAATAACGGGTATTCTCGAGATG 2745

Qy       2892 AGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGA 2951
Db       2746 CAATGGGGCAAAGTTGGGATCGATGATTGGTGGAGAAACGAACAGTTTTGGGTCATTGGA 2805

Qy       2952 GGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTT 3011
Db       2806 GGTGTTTCTGCGCATCTGTTTGCTCTCTTCCAAGGTCTCCTCAAGGTTCTTGCTGGTGTC 2865

Qy       3012 GATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
Db       2866 GACACTAACTTCACAGTCACATCAAAAGCAGCTGATGATGGAGAGTTCTCTGACCTTTAC 2925

Qy       3072 ACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATT 3131
Db       2926 CTCTTCAAATGGACTTCACTTCTCATCCCTCCAATGACTCTACTCATCATAAACGTCATT 2985

Qy       3132 GGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTC 3191
Db       2986 GGAGTCATAGTCGGAGTCTCTGATGCCATCAGCAATGGATACGACTCGTGGGGACCGCTT 3045

Qy       3192 TTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGT 3251
Db       3046 TTCGGAAGACTGTTCTTTGCACTTTGGGTCATCATTCATCTTTACCCGTTCCTTAAAGGT 3105

Qy       3252 TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCT 3311
Db       3106 TTGCTTGGGAAACAAGATAGAATGCCAACCATTATTGTCGTCTGGTCCATCCTCCTGGCC 3165

Qy       3312 TCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGGTCCG 3371
Db       3166 TCGATTCTTACACTTCTTTGGGTCCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCCT 3222

Qy       3372 CTTCTTGAGGAGTGTGGTTTGGATTGC 3398
Db       3223 ATTCTCGAGATCTGTGGTTTAGACTGC 3249
```

RESULT 13
US-09-938-842A-559
; Sequence 559, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 559
; LENGTH: 3255
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-559

Query Match 36.4%; Score 1431.8; DB 10; Length 3255;
Best Local Similarity 66.0%; Pred. No. 0;
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

```
Qy       144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203
Db         1 ATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACAGAAACGAATTCGTTCTCATT 60

Qy       204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db        61 AACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAACTGAGTGGGCAAACATGTCAA 120

Qy       264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db       121 ATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGCTCTTTGTTGCTTGCAACGAA 180

Qy       324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db       181 TGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAACGTAGAGAAGGAAATCAAGCT 240

Qy       384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
Db       241 TGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTAGTCCACGGGTTGATGGAGAT 300

Qy       444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
Db       301 GATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGGATGGACCCT 360

Qy       504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db       361 GAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACACCGGTCGTGGTGGATTGGAT 420

Qy       564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db       421 TCAGCTCCACCTG----------------GCTCTCAGATTCCTCTTTTGACTTATTGTGAT 465

Qy       624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGG 683
Db       466 GAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCGTGCCTCCTTCAACGGGATAT 525

Qy       684 GGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCT 743
Db       526 GGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTGCACCTCCACAGGCGAGATCA 585

Qy       744 ATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGG 803
Db       586 ATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAAGTGTTGCTTGGAAGGACCGT 645

Qy       804 ATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGT 863
Db       646 ATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAGTCATTAAGCATGAAGGAGGA 705

Qy       864 GATGATGGT----------------------------GACGATGCTGATCTACCACTAATG 896
Db       706 AACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATGATCCTGACATGCCTATGATG 765

Qy       897 GATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCA 956
Db       766 GATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTCGTTCAAGCAGAATAAATCCT 825

Qy       957 TATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTG 1016
Db       826 TACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTCTTTTCTTTCATTATAGAATT 885

Qy      1017 ATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGG 1076
Db       886 CTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGTCAGTTATATGCGAGATATGG 945

Qy      1077 TTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACT 1136
Db       946 TTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGTATCCTATAGAACGTGAAACA 1005

Qy      1137 TACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCA 1196
Db       1006 TACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAAAACCGTCAGGATTAGCACCT 1065

Qy      1197 ATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAAT 1256
Db       1066 GTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGCCACCCTTGATTACAGCAAAC 1125

Qy      1257 ACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCT 1316
Db       1126 ACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATAAGGTTGCGTGTTATGTATCA 1185

Qy      1317 GATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAG 1376
Db       1186 GACGATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTGATACAGCTGAGTTTGCTAGA 1245

Qy      1377 AAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTC 1436
Db       1246 AAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCACGAGCTCCTGAGTGGTATTTT 1305

Qy      1437 CAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGA 1496
Db       1306 TCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTGCTTTTGTCAGGGAACGTCGT 1365

Qy      1497 GCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCC 1556
Db       1366 GCTATGAAGAGAGATTATGAGGAGTTTAAAGTGAAGATAAATGCACTGGTTGCTACTGCA 1425

Qy      1557 CAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAAT 1616
Db       1426 CAGAAAGTGCCTGAGGAAGGTTGGACTATGCAAGATGGAACTCCTTGGCCTGGAAACAAC 1485

Qy      1617 GTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGT 1676
Db       1486 GTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTCATAGTGGAGTTCGTGATACG 1545

Qy      1677 GAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAAC 1736
Db       1546 GATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTGAGAAGCGGCCTGGATTTGAT 1605

Qy      1737 CATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAAT 1796
Db       1606 CACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAGTCTCTGCTGTTCTATCAAAC 1665

Qy      1797 GCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAG 1856
Db       1666 GCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCAACAACAGCAAAGCAATTAGA 1725

Qy      1857 GAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTC 1916
Db       1726 GAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGAAAGTTTGTTATGTTCAGTTT 1785

Qy      1917 CCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTT 1976
Db       1786 CCGCAGAGATTTGATGGGATTGATAGACATGATAGATACTCAAACCGTAACGTTGTGTTC 1845

Qy      1977 TTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGA 2036
Db       1846 TTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGACCGATATATGTCGGGACAGGT 1905

Qy      2037 TGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCA 2096
Db       1906 TGTGTGTTTAGAAGACAGGCTCTTTATGGTTTTGATGCACCAAAGAAGAAGAAACCACCA 1965

Qy      2097 TCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGG 2156
Db       1966 GGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGTGTTGTGGGTTGAGAAAGAAG 2025
```

```
Qy       2157 AAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAA 2216
Db       2026 AGTAAAACGAAAGCCAAAGATAAGAAAACTAACACTAAAGAGACTTCA---------- 2073

Qy       2217 GAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCT 2276
Db       2074 ------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAGGTGTTATCGTCCCAGTGTCA 2127

Qy       2277 GAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCT 2336
Db       2128 AATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGGAGAAGAAGTTTGGACAATCT 2187

Qy       2337 TCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCT 2396
Db       2188 CCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAGTTCCCCGTAACGCAAGCCCC 2247

Qy       2397 GCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGAC 2456
Db       2248 GCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGAA 2307

Qy       2457 TGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTC 2516
Db       2308 TGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTGAAGATATCCTGACGGGTTTC 2367

Qy       2517 AAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAA 2576
Db       2368 AAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGCCTAAGCGTGCAGCTTTTAAA 2427

Qy       2577 GGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGG 2636
Db       2428 GGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAGTTCTACGTTGGGCTCTTGGC 2487

Qy       2637 TCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTG 2696
Db       2488 TCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGTATGGTTATGGTGGTGGTTTA 2547

Qy       2697 AAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCC 2756
Db       2548 AAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCTATCCTTGGACTTCACTTCCA 2607

Qy       2757 CTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCA 2816
Db       2608 TTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCACAGGAAAATTCATCGTCCCT 2667

Qy       2817 GAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACG 2876
Db       2668 GAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGTTCATATCCATAGCAGTAACT 2727

Qy       2877 AGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAG 2936
Db       2728 GGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATGATTGGTGGAGAAACGAGCAG 2787

Qy       2937 TTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAG 2996
Db       2788 TTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTCTGTTTCAAGGTTTGCTCAAA 2847

Qy       2997 GTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAG 3056
Db       2848 GTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAAAAGCAGCAGACGATGGAGCT 2907

Qy       3057 TTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTT 3116
Db       2908 TTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGATTCCTCCGACAACACTTCTG 2967

Qy       3117 CTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAG 3176
Db       2968 ATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATGCCATTAGCAATGGCTATGAC 3027

Qy       3177 TCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTAT 3236
Db       3028 TCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTTGGGTCATTGTTCATTTATAC 3087

Qy       3237 CCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGG 3296

Db       3088 CCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGCCTACGATTATTGTGGTCTGG 3147

Qy       3297 TCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCG 3356
Db       3148 TCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCAGAGTTAACCCGTTTGTGGCT 3207

Qy       3357 AAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTG 3397
Db       3208 A---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATTG 3245
```

RESULT 14
US-10-229-193-7
; Sequence 7, Application US/10229193
; Publication No. US20030126643A1
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of Cellulose and/or Beta-1.4-Glucan
; FILE REFERENCE: 96-98A
; CURRENT APPLICATION NUMBER: US/10/229,193
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/221,013
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3828
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (239)..(3490)
US-10-229-193-7

Query Match 36.3%; Score 1429.4; DB 15; Length 3828;
Best Local Similarity 65.8%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;

```
Qy        124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
Db        219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy        184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
Db        279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy        244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
Db        339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy        304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
Db        399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy        364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
Db        459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy        424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483
Db        519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy        484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543
Db        579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638
```

```
Qy       544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603
Db       639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT---------------CAGATTC 683

Qy       604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663
Db       684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743

Qy       664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723
Db       744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803

Qy       724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783
Db       804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863

Qy       784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843
Db       864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923

Qy       844 CGGGGAATGATGGTGGTGGTGATGATGGT---------------------------GACG 876
Db       924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983

Qy       877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936
Db       984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043

Qy       937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996
Db      1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103

Qy       997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056
Db      1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163

Qy      1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116
Db      1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223

Qy      1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176
Db      1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283

Qy      1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236
Db      1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343

Qy      1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296
Db      1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403

Qy      1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356
Db      1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463

Qy      1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416
Db      1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523

Qy      1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476
Db      1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583

Qy      1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536
Db      1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643

Qy      1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596
Db      1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703

Qy      1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656
```

```
Db      1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763

Qy      1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716
Db      1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823

Qy      1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776
Db      1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883

Qy      1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836
Db      1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943

Qy      1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896
Db      1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003

Qy      1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956
Db      2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063

Qy      1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016
Db      2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123

Qy      2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076
Db      2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183

Qy      2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136
Db      2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243

Qy      2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196
Db      2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303

Qy      2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256
Db      2304 AGACTTCA-----------------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345

Qy      2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316
Db      2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405

Qy      2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376
Db      2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465

Qy      2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436
Db      2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525

Qy      2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496
Db      2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585

Qy      2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556
Db      2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645

Qy      2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db      2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705

Qy      2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db      2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765

Qy      2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
```

```
Db       2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825

Qy       2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796

Db       2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885

Qy       2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856

Db       2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945

Qy       2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916

Db       2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005

Qy       2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976

Db       3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036

Db       3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096

Db       3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156

Db       3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216

Db       3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276

Db       3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336

Db       3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396

Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397

Db       3483 G 3483
```

RESULT 15
US-10-209-059-45
; Sequence 45, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-45

Query Match 35.0%; Score 1378.4; DB 13; Length 3725;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;

```
Qy        119 CGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCA 178

Db        154 CGGTGTGTGTAGGAGGAGCGCGGAGATGGCGGCCAACAAGGGGATGGTGGCAGGCTCTCA 213

Qy        179 CAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA------TCCCGGGCCGAAGCC 232

Db        214 CAACCGCAACGAGTTCGTCATGATCCGCCACGACGGCGACGCGCCTGTCCCGGCTAAGCC 273

Qy        233 GCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCC 292

Db        274 CACGAAGAGTGCGAATGGGCAGGTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTCAGC 333

Qy        293 CGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTA 352

Db        334 CACTGGTGATGTCTTTGTTGCCTGCAATGAGTGTGCCTTCCCTGTCTGCCGCCCTTGCTA 393

Qy        353 CGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCG 412

Db        394 TGAGTACGAGCGCAAGGAAGGGAACCAATGCTGCCCTCAGTGCAAGACTAGATACAAGAG 453

Qy        413 CCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGA 472

Db        454 ACAGAAAGGTAGCCCTCGAGTTCATGGTGATGATGAGGAGGAAGATGTTGATGACCTGGA 513

Qy        473 CAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGG 532

Db        514 CAATGAATTCAACTATAA----------------------GCAAGGCAATGGGAAGGG 549

Qy        533 CCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAA 592

Db        550 CCCAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCC 609

Qy        593 CCCCAAT---GTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCA 649

Db        610 ACACCATCGGATTCCACGCCTTACAAGTGGACAACAGATATCTGGAGAGATCCCTGATGC 669

Qy        650 GCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTA 709

Db        670 ATCCCCTGACCGTCATTCTATCCGCAGT----------------CCAACATCGAGCTA 711

Qy        710 TGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGC 769

Db        712 TGTTGATCCAAGCGTTCCAGTTCCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTC 771

Qy        770 ATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGA 829

Db        772 CTATGGGCTTAATAGTGTTGACTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831

Qy        830 GAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTG---------------- 873

Db        832 CAAAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGGAC 891

Qy        874 --------ACGATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAA 925

Db        892 TGGCTCAAATGGAGAAGATATGCAAATGGTTGATGATGCACGCCTACCTTTGAGCCGCAT 951

Qy        926 AATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGT 985

Db        952 TGTGCCAATTTCCTCAAACCAGCTCAACCTTTACCGGATAGTAATCATTCTCCGTCTTAT 1011

Qy        986 GGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTT 1045

Db       1012 CATCCTGTGCTTCTTCTTCCAATATCGTATCAGTCATCCAGTGCGTAATGCTTATGGATT 1071

Qy       1046 GTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATT 1105
```

```
Db       1072 GTGGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGCCTTGTCCTGGCTTCTAGATCAGTT 1131
Qy       1106 CCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGA 1165
Db       1132 CCCAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCATTGAGGTATGA 1191
Qy       1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225
Db       1192 TAGAGAGGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTTGTCAGTACAGTGGATCC 1251
Qy       1226 CTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTA 1285
Db       1252 ATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTTGTCCATTCTTGCTGTGGATTA 1311
Qy       1286 TCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGA 1345
Db       1312 CCCTGTTGACAAAGTGTCATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGA 1371
Qy       1346 AGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAA 1405
Db       1372 GTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAAGAAGCACAA 1431
Qy       1406 TATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAA 1465
Db       1432 TATTGAACCAAGAGCTCCAGAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAA 1491
Qy       1466 GGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAA 1525
Db       1492 AATTCAACCTTCATTTGTTAAGGAAAGACGAGCAATGAAGAGAGAGTATGAAGAATTCAA 1551
Qy       1526 GGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAAT 1585
Db       1552 AATAAGAATCAATGCCCTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCAT 1611
Qy       1586 GCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGT 1645
Db       1612 GGCTGATGGAACTGCTTGGCCTGGGAATAACCCTAGGGACCATCCTGGCATGATTCAGGT 1671
Qy       1646 CTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTA 1705
Db       1672 GTTCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAAATGAATTACCACGTCTTGTCTA 1731
Qy       1706 TGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGC 1765
Db       1732 TGTCTCTCGTGAAAAGAGACCAGGCTTTCAGCATCACAAGAAGGCTGGTGCAATGAATGC 1791
Qy       1766 ATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGA 1825
Db       1792 ACTGATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGTGGATTGTGA 1851
Qy       1826 TCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTT 1885
Db       1852 CCATTACTTCAATAGCAGCAAAGCTCTTAGAGAAGCAATGTGCTTCATGATGGATCCAGC 1911
Qy       1886 ACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCA 1945
Db       1912 TCTAGGAAGGAAAACTTGTTATGTACAATTTCCACAAAGATTTGATGGCATTGACTTGCA 1971
Qy       1946 TGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGG 2005
Db       1972 CGATCGATATGCTAATAGGAACATAGTCTTCTTTGATATCAACATGAAAGGTCTAGATGG 2031
Qy       2006 TATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGG 2065
Db       2032 CATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGG 2091
Qy       2066 TTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTG 2125
Db       2092 ATATGAT-----------------CCTGTTTTGACTGAAGCTGATCTGGAACCTAACATT 2134
Qy       2126 GTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAAC 2185
```

```
Db       2135 GTTGTTAAG----AGCTGCTGTGGTAGAAGGAAGAGAAAGAACAAGAGTTATATGGATAG 2190
Qy       2186 AGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCT 2245
Db       2191 TCAAAGCCGTA---------TTATGAAGAGAACAGAATCTTCAGCTCCCATCTTTAACAT 2241
Qy       2246 TGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCA 2305
Db       2242 GGAAGACATCGAGGAGGGTATTGAAGGTTATGAGGATGAAAGGTCAGTGCTTATGTCCCA 2301
Qy       2306 ACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGA 2365
Db       2302 GAGGAAATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTCATTGCATCCACCTTTATGAC 2361
Qy       2366 GAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGT 2425
Db       2362 TCAAGGTGGCATACCACCTTCAACAAACCCAGCTTCTCTACTGAAGGAAGCTATCCATGT 2421
Qy       2426 CATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGG 2485
Db       2422 TATCAGCTGTGGGTACGAGGACAAAACTGAATGGGGAAAAGAGATTGGCTGGATCTATGG 2481
Qy       2486 ATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAAT 2545
Db       2482 TTCAGTTACAGAGGATATTCTGACTGGGTTTAAAATGCATGCAAGAGGCTGGCAATCAAT 2541
Qy       2546 TTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCG 2605
Db       2542 CTACTGCATGCCACCACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATCG 2601
Qy       2606 TCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTG 2665
Db       2602 TCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAGTGGAAATTCTGCTTAGCAGACATTG 2661
Qy       2666 CCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAA 2725
Db       2662 TCCTATATGGTATGGCTACAATGGGCGATTGAAGCTTTTGGAGAGGCTGGCTTACATTAA 2721
Qy       2726 CTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCAT 2785
Db       2722 CACCATTGTTTATCCAATCACATCTGTTCCGCTTATCGCCTATTGTGTGCTTCCTGCTAT 2781
Qy       2786 CTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTT 2845
Db       2782 CTGTCTTCTTACCAATAAATTTATCATTCCTGAGATTAGTAATTATGCTGGAATGTTCTT 2841
Qy       2846 CATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGT 2905
Db       2842 CATTCTTCTTTTTGCCTCCATTTTCGCAACTGGTATATTGGAGCTCAGATGGAGTGGTGT 2901
Qy       2906 TGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACA 2965
Db       2902 TGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTTGGGTTATTGGTGGCACCTCTGCCCA 2961
Qy       2966 CCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCAC 3025
Db       2962 TCTCTTCGCGGTGTTCCAGGGTCTGCTGAAAGTGTTGGCTGGGATTGATACCAACTTCAC 3021
Qy       3026 CGTGACATCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATG 3082
Db       3022 AGTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTG 3081
Qy       3083 GACTACCTTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGC 3142
Db       3082 GACCAGTTTGCTCATCCCTCCGACCACTGTTCTTGTCATTAACCTGGTCGGAATGGTGGC 3141
Qy       3143 TGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCT 3202
Db       3142 AGGAATTTCGTATGCCATTAACAGCGGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCT 3201
Qy       3203 ATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAG 3262
Db       3202 GTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCCTCAAGGGTCTCATGGGCAG 3261
```

```
Qy       3263 GCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTC 3322
              ||| ||| | || ||||| || |||||||| ||||||||| || || || || ||||||||
Db       3262 GCAGAACCGCACGCCAACAATCGTCATCGTTTGGTCCATCCTCCTTGCGTCTATCTTCTC 3321

Qy       3323 GCTCCTTTGGGTTCGGATTGATCCTTTCCT 3352
                | || |||||   ||| ||||||||| |
Db       3322 CTTGCTGTGGGTGAAGATCGATCCTTTCAT 3351
```

Search completed: December 16, 2003, 04:38:54
Job time : 1289 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 18:07:06 ; Search time 210 Seconds
(without alignments)
8272.781 Million cell updates/sec

Title: US-09-720-383C-9
Perfect score: 3936
Sequence: 1 cttctccctcgtcggtgcgg.....ctgagttggttttgtgaaaa 3936

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1429.4	36.3	3828	4	US-09-221-013A-7	Sequence 7, Appli
2	1330.8	33.8	3673	4	US-09-221-013A-11	Sequence 11, Appl
3	1326.6	33.7	3603	4	US-09-221-013A-5	Sequence 5, Appli
4	1269.2	32.2	3614	4	US-09-221-013A-9	Sequence 9, Appli
5	1134.8	28.8	3328	3	US-08-960-048-1	Sequence 1, Appli
6	1134.8	28.8	3328	4	US-09-838-586-1	Sequence 1, Appli
7	884.6	22.5	2248	4	US-09-221-013A-1	Sequence 1, Appli
8	664.2	16.9	1741	4	US-09-221-013A-13	Sequence 13, Appl
9	397.8	10.1	8411	4	US-09-221-013A-3	Sequence 3, Appli
10	316.8	8.0	5009	4	US-09-221-013A-4	Sequence 4, Appli
11	268	6.8	268	4	US-09-313-294A-849	Sequence 849, App
12	213	5.4	317	4	US-09-313-294A-7123	Sequence 7123, Ap
13	164.2	4.2	278	4	US-09-313-294A-4015	Sequence 4015, Ap
14	159	4.0	281	4	US-09-313-294A-6485	Sequence 6485, Ap
15	145	3.7	300	4	US-09-313-294A-4753	Sequence 4753, Ap
16	144.8	3.7	291	4	US-09-313-294A-5519	Sequence 5519, Ap
17	135.8	3.5	286	4	US-09-313-294A-4696	Sequence 4696, Ap
18	130.6	3.3	287	4	US-09-313-294A-5274	Sequence 5274, Ap
19	99.2	2.5	283	4	US-09-313-294A-1981	Sequence 1981, Ap
20	90.8	2.3	284	4	US-09-313-294A-3993	Sequence 3993, Ap
21	77.6	2.0	159	4	US-09-313-294A-1693	Sequence 1693, Ap
22	73	1.9	250	4	US-09-313-294A-582	Sequence 582, App
c 23	59.8	1.5	1926	4	US-09-249-585A-4	Sequence 4, Appli
c 24	59.8	1.5	1931	2	US-09-130-114-2	Sequence 2, Appli
25	59.6	1.5	390	3	US-09-197-649-7	Sequence 7, Appli
26	56	1.4	44377	2	US-08-804-227C-7	Sequence 7, Appli
27	56	1.4	44377	2	US-08-804-198-1	Sequence 1, Appli
c 28	55.6	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 29	55.6	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 30	54.4	1.4	1080	4	US-09-252-991A-8705	Sequence 8705, Ap
31	54.4	1.4	1233	4	US-09-252-991A-9076	Sequence 9076, Ap
32	54.4	1.4	1293	4	US-09-252-991A-9249	Sequence 9249, Ap
33	53.8	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
34	53.8	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
35	52.4	1.3	4612	3	US-08-960-048-2	Sequence 2, Appli
36	52.4	1.3	4612	4	US-09-838-586-2	Sequence 2, Appli
c 37	51.6	1.3	714	4	US-09-252-991A-15925	Sequence 15925, A
38	51.6	1.3	780	4	US-09-252-991A-15799	Sequence 15799, A
39	51.6	1.3	2055	4	US-09-252-991A-15771	Sequence 15771, A
40	51.2	1.3	2211	4	US-09-252-991A-6288	Sequence 6288, Ap
c 41	51.2	1.3	2328	4	US-09-252-991A-6020	Sequence 6020, Ap
42	49.6	1.3	2249	3	US-08-814-052-19	Sequence 19, Appl
43	49.6	1.3	2279	3	US-08-814-052-17	Sequence 17, Appl
44	49.6	1.3	2300	3	US-08-814-052-18	Sequence 18, Appl
45	49.6	1.3	3183	2	US-08-939-218A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-221-013A-7
; Sequence 7, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3828
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (239)..(3490)
US-09-221-013A-7

Query Match 36.3%; Score 1429.4; DB 4; Length 3828;
Best Local Similarity 65.8%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;

```
Qy      124 GAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACC 183
            ||   || |  || ||    ||| |  |  | |   ||||  | || ||||| ||||||
Db      219 GATTCGGTGTAGAAGACATCATGAATACTGGTGGTCGGCTCATTGCTGGCTCTCACAACA 278

Qy      184 GCAACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGC 243
            | |||||  ||||  ||||  |  ||| |    | || |       |    | || |
Db      279 GAAACGAATTCGTTCTCATTAACGCCGATGAGAGTGCCAGAATACGATCAGTACAAGAAC 338

Qy      244 AGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACC 303
              ||  |||||    || || || || || || || ||  |||   |  |     || | || |
Db      339 TGAGTGGGCAAACATGTCAAATCTGTGGAGATGAAATCGAATTAACGGTTAGCAGTGAGC 398

Qy      304 CCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGC 363
              ||| || || |||||||| ||||| ||||| || ||  |    |||| || || || |
Db      399 TCTTTGTTGCTTGCAACGAATGCGCATTCCCGGTTTGTAGACCATGCTATGAGTATGAAC 458

Qy      364 GCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCT 423
            |  | || || |  ||    || || ||||||| ||||||||||  |  | || ||
```

```
Db        459 GTAGAGAAGGAAATCAAGCTTGTCCTCAGTGCAAAACTCGATACAAAAGGATTAAAGGTA 518

Qy        424 GCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCA 483

Db        519 GTCCACGGGTTGATGGAGATGATGAAGAAGAAGAAGACATTGATGATCTTGAGTATGAGT 578

Qy        484 ACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543

Db        579 TTGATCATGGGATGGACCCTGAACATGCCGCTGAAGCCGCACTCTCTTCACGCCTTAACA 638

Qy        544 ACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTC 603

Db        639 CCGGTCGTGGTGGATTGGATTCAGCTCCACCTGGCTCT---------------CAGATTC 683

Qy        604 CACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGC 663

Db        684 CTCTTTTGACTTATTGTGATGAAGATGCTGATATGTATTCTGATCGTCATGCTCTTATCG 743

Qy        664 CTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCT 723

Db        744 TGCCTCCTTCAACGGGATATGGGAATCGCGTCTATCCTGCACCGTTTACAGATTCTTCTG 803

Qy        724 TACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTA 783

Db        804 CACCTCCACAGGCGAGATCAATGGTTCCTCAGAAAGATATTGCGGAATATGGTTATGGAA 863

Qy        784 GTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGA 843

Db        864 GTGTTGCTTGGAAGGACCGTATGGAAGTTTGGAAGAGACGACAAGGCGAAAAGCTTCAAG 923

Qy        844 CGGGGAATGATGGTGGTGGTGATGATGGT--------------------------GACG 876

Db        924 TCATTAAGCATGAAGGAGGAAACAATGGTCGAGGTTCCAATGATGACGACGAACTAGATG 983

Qy        877 ATGCTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTC 936

Db        984 ATCCTGACATGCCTATGATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTC 1043

Qy        937 CATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGT 996

Db       1044 GTTCAAGCAGAATAAATCCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTC 1103

Qy        997 TCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATAT 1056

Db       1104 TTTTCTTTCATTATAGAATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGT 1163

Qy       1057 CTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGT 1116

Db       1164 CAGTTATATGCGAAATATGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGT 1223

Qy       1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCC 1176

Db       1224 ATCCTATAGAACGTGAAACATACCTCGATAGACTCTCTCTCAGGTACGAGAAGGAAGGAA 1283

Qy       1177 AGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC 1236

Db       1284 AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC 1343

Qy       1237 CTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATA 1296

Db       1344 CCCCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTATCCTGTGGATA 1403

Qy       1297 AGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTG 1356

Db       1404 AGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG 1463

Qy       1357 AAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTC 1416

Db       1464 ATACAGCTGATTTTGCTACAAAATGGGTTCCTTTTTGTAAGAAGTTTAATATCGAGCCAC 1523

Qy       1417 GCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAA 1476

Db       1524 GAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCATCCTG 1583

Qy       1477 ACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCA 1536

Db       1584 CTTTTGTCAGGGAACGTCGTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA 1643

Qy       1537 ATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAA 1596

Db       1644 ATGCACTGGTTGCTACTGCACAGAAAGTGCCTGAGGAACGTTGGACTATGCAAGATGGAA 1703

Qy       1597 CCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCC 1656

Db       1704 CTCCTTGGCCTGGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTGTTCTTGGGTC 1763

Qy       1657 AAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAG 1716

Db       1764 ATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTCTAGTGTATGTTTCTCGTG 1823

Qy       1717 AGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG 1776

Db       1824 AGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTATGAATTCCTTGATCCGAG 1883

Qy       1777 TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCA 1836

Db       1884 TCTCTGCTGTTCTATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCA 1943

Qy       1837 ACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGA 1896

Db       1944 ACAACAGCAAAGCAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGA 2003

Qy       1897 AGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATG 1956

Db       2004 AAGTTTGTTATGTTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACT 2063

Qy       1957 CTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTC 2016

Db       2064 CAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAAGGTCTTGATGGGATACAAGGAC 2123

Qy       2017 CAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCC 2076

Db       2124 CGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCTCTTTATGGTTTTGATGCAC 2183

Qy       2077 CCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCT 2136

Db       2184 CAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGGCCTAAATGGTGTTGTTTGT 2243

Qy       2137 GTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAA 2196

Db       2244 GTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCACAGATAAGAAAACTAACACTAAAG 2303

Qy       2197 AGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTG 2256

Db       2304 AGACTTCA------------------AAGCAGATTCATGCGCTAGAGAATGTCGACGAAG 2345

Qy       2257 ACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAG 2316

Db       2346 GTGTTATCGTCCCAGTGTCAAATGTTGAGAAGAGATCTGAAGCAACACAATTGAAATTGG 2405

Qy       2317 AAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAA 2376

Db       2406 AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG 2465

Qy       2377 CCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTG 2436

Db       2466 TTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGCG 2525

Qy       2437 GTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAG 2496

Db       2526 GGTACCAAGATAAAACCGAATGGGGAAAAGAGATCGGGTGGATTTATGGATCGGTGACTG 2585

Qy       2497 AAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATAC 2556

Db       2586 AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC 2645
```

```
Qy       2557 CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGG 2616
Db       2646 CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAG 2705

Qy       2617 TGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGT 2676
Db       2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGT 2765

Qy       2677 ATGGGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGT 2736
Db       2766 ATGGTTATGGTGGTGGTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCT 2825

Qy       2737 ATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGA 2796
Db       2826 ATCCTTGGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCA 2885

Qy       2797 CAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTT 2856
Db       2886 CAGGAAAATTCATCGTCCCTGAGATAAGCAACTACGCAGGTATACTCTTCATGCTCATGT 2945

Qy       2857 TTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATG 2916
Db       2946 TCATATCCATAGCAGTAACTGGAATCCTCGAAATGCAATGGGGAGGTGTCGGAATCGATG 3005

Qy       2917 ACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTG 2976
Db       3006 ATTGGTGGAGAAACGAGCAGTTTTGGGTAATCGGAGGGGCCTCCTCGCATCTATTTGCTC 3065

Qy       2977 TGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAA 3036
Db       3066 TGTTTCAAGGTTTGCTCAAAGTTCTAGCCGGAGTTAACACGAATTTCACAGTCACTTCAA 3125

Qy       3037 AGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db       3126 AAGCAGCAGACGATGGAGCTTTCTCTGAGCTTTACATCTTCAAGTGGACAACTTTGTTGA 3185

Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db       3186 TTCCTCCGACAACACTTCTGATCATTAACATCATTGGAGTTATTGTCGGCGTTTCTGATG 3245

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db       3246 CCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGGAGACTTTTCTTCGCTCTTT 3305

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db       3306 GGGTCATTGTTCATTTATACCCATTCCTCAAGGGAATGCTTGGGAAGCAAGACAAAATGC 3365

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db       3366 CTACGATTATTGTGGTCTGGTCTATTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA 3425

Qy       3337 GGATTGATCCTTTCCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATT 3396
Db       3426 GAATTAACCCGTTTGTGGCTA---AAGGGGGACCAGTGTTGGAGATCTGTGGTCTGAATT 3482

Qy       3397 G 3397
Db       3483 G 3483
```

RESULT 2
US-09-221-013A-11
; Sequence 11, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3673
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)..(3313)
US-09-221-013A-11

Query Match 33.8%; Score 1330.8; DB 4; Length 3673;
Best Local Similarity 64.9%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 1067; Indels 75; Gaps 7;

```
Qy        127 GTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCA 186
Db         54 GTGTGTCGGTGGCTGCGATGGAGGCCAGTGCCGGCTTGGTTGCTGGATCCTACCGGAGAA 113

Qy        187 ACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGA 246
Db        114 ACGAGCTCGTTCGGATCCGACATGAATCTGATGGCGGGACCAAACCTTTGAAGAATATGA 173

Qy        247 ACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCT 306
Db        174 ATGGCCAGATATGTCAGATCTGTGGTGATGATGTTGGACTCGCTGAAACTGGAGATGTCT 233

Qy        307 TCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCC 366
Db        234 TTGTCGCGTGTAATGAATGTGCCTTCCCTGTGTGTCGGCCTTGCTATGAGTACGAGAGGA 293

Qy        367 GGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCC 426
Db        294 AAGATGGAACTCAGTGTTGCCCTCAATGCAAGACTAGATTCAGACGACACAGGGGGAGTC 353

Qy        427 AACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACT 486
Db        354 CTCGTGTTGAAGGAGATGAAGATGAGGATGATGTTGATGATATCGAGAATGAGTTCA--- 410

Qy        487 GGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACG 546
Db        411 ---------ATTACGCCCAGGGAGCTAACAAGGCGAGACACCAACGCCA---TGGCGAAG 458

Qy        547 GCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCAC 606
Db        459 AGTTTTCTTCTTCCTCTAGACATGAATCTCAACCAATTCCTCTTCTCACCCATGGCCATA 518

Qy        607 TCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTT 666
Db        519 CGGTTTCTGGAGAGATTCGCACGCCTGATACACAATCTGTGCGAACTACATCAGGTCCTT 578

Qy        667 CTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTAC 726
Db        579 ------TGGGTCCTTCTGACAGGAATGCTATTTCATCTCCATATATTGATCCACGGCAAC 632

Qy        727 CTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTG 786
Db        633 CTGTCCCTGTAAGAATCGTGGACCCGTCAAAAGACTTGAACTCTTATGGGCTTGGTAATG 692

Qy        787 TTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGGATG---------- 836
Db        693 TTGACTGGAAAGAAAGAGTTGAAGGCTGGAAGCTGAAGCAGGAGAAAAATATGTTACAGA 752

Qy        837 ----------------CACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATG 879
Db        753 TGACTGGTAAATACCATGAAGGGAAAGGAGGAGAAATTGAAGGGACTGGTTCCAATGGCG 812

Qy        880 CTGATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCAT 939
```

```
Db       813 AAGAACTCCAAATGGCTGATGATACACGTCTTCCTATGAGTCGTGTGGTGCCTATCCCAT 872
Qy       940 CAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCT 999
Db       873 CTTCTCGCCTAACCCCTTATCGGGTTGTGATTATTCTCCGGCTTATCATCTTGTGTTTCT 932
Qy      1000 TCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTG 1059
Db       933 TCTTGCAATATCGTACAACTCACCCTGTGAAAAATGCATATCCTTTGTGGTTGACCTCGG 992
Qy      1060 TTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCC 1119
Db       993 TTATCTGTGAGATCTGGTTTGCATTTTCTTGGCTTCTTGATCAGTTTCCCAAATGGTACC 1052
Qy      1120 CTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGC 1179
Db      1053 CCATTAACAGGGAGACTTATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAAC 1112
Qy      1180 CATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTC 1239
Db      1113 CATCACAGCTCGTTCCTGTTGATGTGTTTGTTAGTACAGTGGACCCATTGAAAGAGCCTC 1172
Qy      1240 CTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGG 1299
Db      1173 CCCTTGTTACAGCAAACACAGTTCTCTCGATTCTTTCTGTGGACTACCCGGTAGATAAAG 1232
Qy      1300 TTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAA 1359
Db      1233 TAGCCTGTTATGTTTCAGATGATGGTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAA 1292
Qy      1360 CATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCG 1419
Db      1293 CCGCTGAGTTTGCAAAGAAATGGGTACCATTTTGCAAGAAATTCAACATTGAACCTAGGG 1352
Qy      1420 CTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACT 1479
Db      1353 CCCCTGAATTCTATTTTGCCCAGAAGATAGATTACTTGAAGGACAAGATCCAACCGTCTT 1412
Qy      1480 TTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATG 1539
Db      1413 TTGTTAAAGAGCGACGAGCTATGAAGAGAGAGTATGAAGAGTTTAAAGTGAGGATAAATG 1472
Qy      1540 CCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCC 1599
Db      1473 CTCTTGTTGCCAAAGCACAGAAAATCCCTGAAGAAGGCTGGACAATGCAGGATGGTACTC 1532
Qy      1600 CCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAA 1659
Db      1533 CCTGGCCTGGTAACAACACTAGAGATCATCCTGGAATGATACAGGTGTTCTTAGGCCATA 1592
Qy      1660 GCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGA 1719
Db      1593 GTGGGGGTCTGGATACCGATGGAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAA 1652
Qy      1720 AACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCT 1779
Db      1653 AGCGGCCTGGATTTCAACACCACAAAAAGGCTGGAGCTATGAATGCATTGATCCGTGTAT 1712
Qy      1780 CTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACA 1839
Db      1713 CTGTTGTTCTTACCAATGGAGCATATCTTTTGAACGTGGATTGTGATCATTACTTTAATA 1772
Qy      1840 ACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGG 1899
Db      1773 ACAGTAAGGCTATTAAAGAAGCTATGTGTTTCATGATGGACCCGGCTATTGGAAAGAAGT 1832
Qy      1900 TTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTA 1959
Db      1833 GCTGCTATGTCCAGTTCCCTCAACGTTTTGACGGTATTGATTTGCACGATCGATATGCCA 1892
Qy      1960 ACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAA 2019
```

```
Db      1893 ACAGGAATATAGTCTTTTTCGATATTAACATGAAGGGGTTGGATGGTATCCAGGGTCCAG 1952
Qy      2020 TTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCA 2079
Db      1953 TATATGTGGGTACTGGTTGTTGTTTTAATAGGCAGGCTCTATATGGGTATGATCCTGTTT 2012
Qy      2080 AAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTT 2139
Db      2013 TGACGGAAGAAGATTTAGAACCAAATATTA----------------------TTGTCAAGA 2051
Qy      2140 GCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGT 2199
Db      2052 GCTGTTGCGGGTCAAGGAAGAAAGGTAAAAGTAGCAA---GAAGTATAACTACGAAAAGA 2108
Qy      2200 TATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACG 2259
Db      2109 GGAGAGGCATCAACAGAAGTGACTCCAATGCTCCACTTTTCAATATGGAGGACATCGATG 2168
Qy      2260 AAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAA 2319
Db      2169 AGGGTTTTGAAGGTTATGATGATGAGAGGTCTATTCTAATGTCCCAGAGGAGTGTAGAGA 2228
Qy      2320 AGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCT 2379
Db      2229 AGCGTTTTGGTCAGTCGCCGGTATTTATTGCGGCAACCTTCATGGAACAAGGCGGCATTC 2288
Qy      2380 TGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTT 2439
Db      2289 CACCAACAACCAATCCCGCTACTCTTCTGAAGGAGGCTATTCATGTTATAAGCTGTGGTT 2348
Qy      2440 ATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAG 2499
Db      2349 ACGAAGACAAGACTGAATGGGGCAAAGAGATTGGTTGGATCTATGGTTCCGTGACGGAAG 2408
Qy      2500 ATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTA 2559
Db      2409 ATATTCTTACTGGGTTCAAGATGCATGCCCGGGGTTGGATATCGATCTACTGCAATCCTC 2468
Qy      2560 AACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGC 2619
Db      2469 CACGCCCTGCGTTCAAGGGATCTGCACCAATCAATCTTTCTGATCGTTTGAACCAAGTTC 2528
Qy      2620 TTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATG 2679
Db      2529 TTCGATGGGCTTTGGGATCTATCGAGATTCTTCTTAGCAGACATTGTCCTATCTGGTATG 2588
Qy      2680 GGTATGGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATC 2739
Db      2589 GTTACCATGGAAGGTTGAGACTTTTGGAGAGGATCGCTTATATCAACACCATCGTCTATC 2648
Qy      2740 CTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAG 2799
Db      2649 CTATTACATCCATCCCTCTTATTGCGTATTGTATTCTTCCCGCTTTTTGTCTCATCACCG 2708
Qy      2800 GGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTA 2859
Db      2709 ACAGATTCATCATACCCGAGATAAGCAACTACGCGAGTATTTGGTTCATTCTACTCTTCA 2768
Qy      2860 TCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACT 2919
Db      2769 TCTCAATTGCTGTGACTGGAATCCTGGAGCTGAGATGGAGCGGTGTGAGCATTGAGGATT 2828
Qy      2920 GGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGT 2979
Db      2829 GGTGGAGGAACGAGCAGTTCTGGGTCATTGGTGGCACATCCGCCCATCTTTTTGCTGTCT 2888
Qy      2980 TCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGG 3039
Db      2889 TCCAAGGTCTACTTAAGGTTCTTGCTGGTATCGACACCAACTTCACCGTTACATCTAAAG 2948
Qy      3040 ---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA 3096
Db      2949 CCACAGACGAAGATGGGGATTTTGCAGAACTCTACATCTTCAAATGGACAGCTCTTCTCA 3008
```

```
Qy       3097 TACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATG 3156
Db       3009 TTCCACCAACCACCGTCCTACTTGTGAACCTCATAGGCATTGTGGCTGGTGTCTCTTATG 3068

Qy       3157 CGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT 3216
Db       3069 CTGTAAACAGTGGCTACCAGTCGTGGGGTCCGCTTTTCGGGAAGCTCTTCTTCGCCTTAT 3128

Qy       3217 GGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACAC 3276
Db       3129 GGGTTATTGCCCATCTCTACCCTTTCTTGAAAGGTCTGTTGGGAAGACAAAACCGAACAC 3188

Qy       3277 CAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTC 3336
Db       3189 CAACCATCGTCATTGTCTGGTCTGTTCTTCTCGCCTCCATCTTCTCGTTGCTTTGGGTCA 3248

Qy       3337 GGATTGATCCTTTCCTTG 3354
Db       3249 GGATCAATCCCTTTGTGG 3266
```

RESULT 3
US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3243)
US-09-221-013A-5

Query Match 33.7%; Score 1326.6; DB 4; Length 3603;
Best Local Similarity 65.0%; Pred. No. 0;
Matches 2107; Conservative 0; Mismatches 1059; Indels 75; Gaps 7;

```
Qy        144 ATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCCCACAACCGCAACGAGCTCGTCGTCATC 203
Db          1 ATGGAGGCCAGTGCCGGCTTGGTTGCTGGATCCTACCGGAGAAACGAGCTCGTTCGGATC 60

Qy        204 CGCCGCGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAG 263
Db         61 CGACATGAATCTGATGGCGGGACCAAACCTTTGAAGAATATGAATGGCCAGATATGTCAG 120

Qy        264 ATTTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAG 323
Db        121 ATCTGTGGTGATGATGTTGGACTCGCTGAAACTGGAGATGTCTTTGTCGCGTGTAATGAA 180

Qy        324 TGCGCCTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAAC 383
Db        181 TGTGCCTTCCCTGTGTGTCGGCCTTGCTATGAGTACGAGAGGAAAGATGGAACTCAGTGT 240

Qy        384 TGCCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443
Db        241 TGCCCTCAATGCAAGACTAGATTCAGACGACACAGGGGGAGTCCTCGTGTTGAAGGAGAT 300

Qy        444 GAGGAGGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503
Db        301 GAAGATGAGGATGATGTTGATGATATCGAGAATGAGTTCA------------ATTACGCC 348

Qy        504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCT 563
Db        349 CAGGGAGCTAACAAGGCGAGACACCAACGCCA---TGGCGAAGAGTTTTCTTCTTCCTCT 405

Qy        564 AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA 623
Db        406 AGACATGAATCTCAACCAATTCCTCTTCTCACCCATGGCCATACGGTTTCTGGAGAGATT 465

Qy        624 ATGGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGG 683
Db        466 CGCACGCCTGATACACAATCTGTGCGAACTACATCAGGTCCTT------TGGGTCCTTCT 519

Qy        684 GGAAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCT 743
Db        520 GACAGGAATGCTATTTCATCTCCATATATTGATCCACGGCAACCTGTCCCTGTAAGAATC 579

Qy        744 ATGGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGG 803
Db        580 GTGGACCCGTCAAAAGACTTGAACTCTTATGGGCTTGGTAATGTTGACTGGAAAGAAAGA 639

Qy        804 ATGGAGAATTGGAAGCAGAGACAAGAGAGGATGC------------------------- 837
Db        640 GTTGAAGGCTGGAAGCTGAAGCAGGAGAAAAATATGTTACAGATGACTGGTAAATACCAT 699

Qy        838 -ACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTACCACTAATG 896
Db        700 GAAGGGAAAGGAGGAGAAATTGAAGGGACTGGTTCCAATGGCGAAGAACTCCAAATGGCT 759

Qy        897 GATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCA 956
Db        760 GATGATACACGTCTTCCTATGAGTCGTGTGGTGCCTATCCCATCTTCTCGCCTAACCCCT 819

Qy        957 TATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTG 1016
Db        820 TATCGGGTTGTGATTATTCTCCGGCTTATCATCTTGTGTTTCTTCTTGCAATATCGTACA 879

Qy       1017 ATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGG 1076
Db        880 ACTCACCCTGTGAAAAATGCATATCCTTTGTGGTTGACCTCGGTTATCTGTGAGATCTGG 939

Qy       1077 TTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACT 1136
Db        940 TTTGCATTTTCTTGGCTTCTTGATCAGTTTCCCAAATGGTACCCCATTAACAGGGAGACT 999

Qy       1137 TACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCA 1196
Db       1000 TATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAACCATCACAGCTCGTTCCT 1059

Qy       1197 ATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAAT 1256
Db       1060 GTTGATGTGTTTGTTAGTACAGTGGACCCATTGAAAGAGCCTCCCCTTGTTACAGCAAAC 1119

Qy       1257 ACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCT 1316
Db       1120 ACAGTTCTCTCGATTCTTTCTGTGGACTACCCGGTAGATAAAGTAGCCTGTTATGTTTCA 1179

Qy       1317 GATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAG 1376
Db       1180 GATGATGGTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAACCGCTGAGTTTGCAAAG 1239

Qy       1377 AAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTC 1436
Db       1240 AAATGGGTACCATTTTGCAAGAAATTCAACATTGAACCTAGGGCCCCTGAATTCTATTTT 1299

Qy       1437 CAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGA 1496
Db       1300 GCCCAGAAGATAGATTACTTGAAGGACAAGATCCAACCGTCTTTTGTTAAAGAGCGACGA 1359
```

```
Qy       1497 GCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCC 1556
             || ||||||||||||||||| || || || ||||| || |||||  | ||||||||||
Db       1360 GCTATGAAGAGAGAGTATGAAGAGTTTAAAGTGAGGATAAATGCTCTTGTTGCCAAAGCA 1419

Qy       1557 CAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAAT 1616
             |||||| | |||||||||||| |||||||||| ||||| || |||||||||||| |||||
Db       1420 CAGAAAATCCCTGAAGAAGGCTGGACAATGCAGGATGGTACTCCCTGGCCTGGTAACAAC 1479

Qy       1617 GTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGT 1676
               | | |||||||||||||||||| ||||| ||| | ||||| || || || || ||
Db       1480 ACTAGAGATCATCCTGGAATGATACAGGTGTTCTTAGGCCATAGTGGGGGTCTGGATACC 1539

Qy       1677 GAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAAC 1736
             || |||||||| |||||  || |  | ||||||||| | || || || || || | | |
Db       1540 GATGGAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAAAGCGGCCTGGATTTCAA 1599

Qy       1737 CATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAAT 1796
             || || || || ||||| |||||||||||||| |||| || |||||||| || || |||
Db       1600 CACCACAAAAAGGCTGGAGCTATGAATGCATTGATCCGTGTATCTGCTGTTCTTACCAAT 1659

Qy       1797 GCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAG 1856
             |  ||||| | || ||| |||||||||||||| ||| | || ||||| |||||||| ||
Db       1660 GGAGCATATCTTTTGAACGTGGATTGTGATCATTACTTTAATAACAGTAAGGCTATTAAA 1719

Qy       1857 GAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTC 1916
             ||||| |||||||| ||||||||||      | |||||||||   |||||||| ||||||
Db       1720 GAAGCTATGTGTTTCATGATGGACCCGGCTATTGGAAAGAAGTGCTGCTATGTCCAGTTC 1779

Qy       1917 CCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTT 1976
             |||||| | ||||| || |||||| || || |||||||| ||| ||||| | |||||||
Db       1780 CCTCAACGTTTTGACGGTATTGATTTGCACGATCGATATGCCAACAGGAATATAGTCTTT 1839

Qy       1977 TTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGA 2036
             || ||||| |||||||| || |||||||||| ||||||||| | ||||| ||||||||
Db       1840 TTCGATATTAACATGAAGGGGTTGGATGGTATCCAGGGTCCAGTATATGTGGGTACTGGT 1899

Qy       2037 TGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCA 2096
             |||   |||| ||||||||  |||||||  |||||| |          || |  |      |
Db       1900 TGTTGTTTTAATAGGCAGGCTCTATATGGGTATGATCCTGTTTTGACGGAAGAAGATTTA 1959

Qy       2097 TCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGG 2156
               |  | |  |                          ||  |    ||||  |  ||    |||
Db       1960 GAACCAAATATTA----------------------TTGTCAAGAGCTGTTGCGGGTCAAGG 1998

Qy       2157 AAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAA 2216
             |||  ||   || | ||  |||    ||      |  | |||||     |    |||| | |
Db       1999 AAGAAAGGTAAAAGTAGCAA---GAAGTATAACTACGAAAAGAGGAGAGGCATCAACAGA 2055

Qy       2217 GAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCT 2276
                ||   | |  | ||    |    | | |  || || || || | |  |||    |
Db       2056 AGTGACTCCAATGCTCCACTTTTCAATATGGAGGACATCGATGAGGGTTTTGAAGGTTAT 2115

Qy       2277 GAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCT 2336
             ||  |||| | | |   | |  |    ||     |  |||| |||   ||||| || ||
Db       2116 GATGATGAGAGGTCTATTCTAATGTCCCAGAGGAGTGTAGAGAAGCGTTTTGGTCAGTCG 2175

Qy       2337 TCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCT 2396
               | || ||| || |  | ||  | | ||  | || || |       |  | | |||
Db       2176 CCGGTATTTATTGCGGCAACCTTCATGGAACAAGGCGGCATTCCACCAACAACCAATCCC 2235

Qy       2397 GCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGAC 2456
             ||| ||||| |||| || ||||| ||||| || || |||||||| |||||||||| ||
Db       2236 GCTACTCTTCTGAAGGAGGCTATTCATGTTATAAGCTGTGGTTACGAAGACAAGACTGAA 2295

Qy       2457 TGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTC 2516
             ||||| |||||||||| |||||||||| || || || |||||||||| ||||| |||
Db       2296 TGGGGCAAAGAGATTGGTTGGATCTATGGTTCCGTGACGGAAGATATTCTTACTGGGTTC 2355

Qy       2517 AAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAA 2576
             |||||||||    |  ||||||   || || |||||||  |||  |||   ||| |||||
Db       2356 AAGATGCATGCCCGGGGTTGGATATCGATCTACTGCAATCCTCCACGCCCTGCGTTCAAG 2415

Qy       2577 GGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGG 2636
```

```
                || ||||||||  | |||||||| ||||||  |  ||||  || |||||  ||||||  | ||
Db       2416 GGATCTGCACCAATCAATCTTTCTGATCGTTTGAACCAAGTTCTTCGATGGGCTTTGGGA 2475

Qy       2637 TCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTG 2696
             |||||  |||||  |  | ||||  |||||  |||  | |||||||||  ||   |||  |  ||
Db       2476 TCTATCGAGATTCTTCTTAGCAGACATTGTCCTATCTGGTATGGTTACCATGGAAGGTTG 2535

Qy       2697 AAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCC 2756
             | | |||||||  ||   |  | || ||||||  |||||||  ||||||    |||||  || ||
Db       2536 AGACTTTTGGAGAGGATCGCTTATATCAACACCATCGTCTATCCTATTACATCCATCCCT 2595

Qy       2757 CTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCA 2816
             ||  | || || ||||  | || ||  | ||| |  | || |  | ||| ||||  ||
Db       2596 CTTATTGCGTATTGTATTCTTCCCGCTTTTTGTCTCATCACCGACAGATTCATCATACCC 2655

Qy       2817 GAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACG 2876
             ||| | |  ||    || ||  | |||||||||    ||| || ||||  |||  ||  ||
Db       2656 GAGATAAGCAACTACGCGAGTATTTGGTTCATTCTACTCTTCATCTCAATTGCTGTGACT 2715

Qy       2877 AGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAG 2936
               | |||||  || |||||||||| |||||  | |||||  || |||||||||||  |||||||
Db       2716 GGAATCCTGGAGCTGAGATGGAGCGGTGTGAGCATTGAGGATTGGTGGAGGAACGAGCAG 2775

Qy       2937 TTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAG 2996
             |||||||||||||| ||    ||| | || || |||||||| ||||| || || || |||
Db       2776 TTCTGGGTCATTGGTGGCACATCCGCCCATCTTTTTGCTGTCTTCCAAGGTCTACTTAAG 2835

Qy       2997 GTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGG---GTGGAGATGATGAG 3053
             ||  | |||||| | || || | |||||||| ||||| || |     |  || |||| |
Db       2836 GTTCTTGCTGGTATCGACACCAACTTCACCGTTACATCTAAAGCCACAGACGAAGATGGG 2895

Qy       3054 GAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTG 3113
             || ||  |||| || || |  |||||||||| |  |  | || || || |||||| |
Db       2896 GATTTTGCAGAACTCTACATCTTCAAATGGACAGCTCTTCTCATTCCACCAACCACCGTC 2955

Qy       3114 CTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATAT 3173
              || ||  ||||| |||| ||  | || ||||| || ||  ||||  | || |  || ||
Db       2956 CTACTTGTGAACCTCATAGGCATTGTGGCTGGTGTCTCTTATGCTGTAAACAGTGGCTAC 3015

Qy       3174 GAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTT 3233
               |||| ||||| || || || |||||||| ||||| || || |||||  |||| ||||||
Db       3016 CAGTCGTGGGGTCCGCTTTTCGGGAAGCTCTTCTTCGCCTTATGGGTTATTGCCCATCTC 3075

Qy       3234 TATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTC 3293
             || || ||  | |||||| || | ||||| |||||| | |||||||| || ||||| |||
Db       3076 TACCCTTTCTTGAAAGGTCTGTTGGGAAGACAAAACCGAACACCAACCATCGTCATTGTC 3135

Qy       3294 TGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTT 3353
             |||||  |||| || || ||||||||| | |||||||  ||||  |||| ||  |
Db       3136 TGGTCTGTTCTTCTCGCCTCCATCTTCTCGTTGCTTTGGGTCAGGATCAATCCCTTTGTG 3195

Qy       3354 G 3354
              |
Db       3196 G 3196
```

RESULT 4
US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; PILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27

```
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 3614
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (217)..(3411)
US-09-221-013A-9

  Query Match             32.2%;  Score 1269.2;  DB 4;  Length 3614;
  Best Local Similarity   63.8%;  Pred. No. 0;
  Matches 2051;  Conservative    0;  Mismatches 1108;  Indels   57;  Gaps    6;

Qy        209 CGACGGCGATCCCGGGCCGAAGCCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTG 268
Db        225 CGAAGGAGAAACCGCGGGAAAGCCGATGAAGAACATTGTTCCGCAGACTTGCCAGATCTG 284

Qy        269 CGGCGACGACGTCGGCCTTGCCCCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGC 328
Db        285 TAGTGACAATGTTGGCAAGACTGTTGATGGAGATCGTTTTGTGGCTTGTGATATTTGTTC 344

Qy        329 CTTCCCCGTCTGCCGGGACTGCTACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCC 388
Db        345 ATTCCCAGTTTGTCGGCCTTGCTACGAGTATGAGAGGAAAGATGGGAATCAATCTTGTCC 404

Qy        389 CCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGA 448
Db        405 TCAGTGCAAAACCAGATACAAGAGGCTCAAAGGTAGTCCTGCTATTCCTGGTGATAAAGA 464

Qy        449 GGAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATGACTCGCAGTC 508
Db        465 CGAGGATGGCTTAGCTGATGAAGGTACTG---TTGAGTTCAACTACCCTCAGAAGGAGAA 521

Qy        509 TGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGG 568
Db        522 AATTTCAGAGCGGATGCTTGGTTGGCATCTTACTCGTGGGAAGGGAGAGGAAATGGGGGA 581

Qy        569 CGCGCCACAAGCTTTCCAGCTCAACCCCA---ATGTTCCACTCCTCACCAACGGGCAAAT 625
Db        582 ACCCCAGTATGATAAAGAGGTCTCTCACAATCATCTTCCTCGTCTCACGAGCAGACAAGA 641

Qy        626 GGTGGATGACATCCCACCGGAGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGG 685
Db        642 TACTTCAGGAGAGTTTTCTGCTGCCTCACCTGAACGCCTCTCTGTATCTTCTACTATCGC 701

Qy        686 AAAGAGGATACATCCCCTTCCTTATGCGGATCCCAGCTTACCTGTGCAACCCAGGTCTAT 745
Db        702 TGGGGGAAAGCGCCTTCCCTATTCATCAGATGTCAATCAATCACCAAATAGAAGGATTGT 761

Qy        746 GGACCCATCCAAGGATCTTGCTGCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGAT 805
Db        762 GGATCCTGTTGGACTCGGGAATGTAGCTTGGAAGGAGAGAGTTGATGGCTGGAAAATGAA 821

Qy        806 GGAGAATTGGAAGCAGAGACAAGAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGA 865
Db        822 GCAAGAGAAGAATACTGGTCCTGTCAGCACGCAGGCTGCTTCTGAAAGAGGTGGAGTAGA 881

Qy        866 TGATGGTGACGATGCTGATCTACCA---------------CTAATGGATGAAGCAAGACA 910
Db        882 TATTGATGCCAGCACAGATATCCTAGCAGATGAGGCTCTGCTGAATGACGAAGCGAGGCA 941

Qy        911 ACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTAT 970
Db        942 GCTTCTGTCAAGGAAAGTTTCAATTCCTTCATCACGGATCAATCCTTACAGAATGGTTAT 1001

Qy        971 CATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAA 1030
Db       1002 TATGCTGCGGCTTGTTATCCTTTGTCTCTTCTTGCATTACCGTATAACAAACCCAGTGCC 1061

Qy       1031 TGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTG 1090
Db       1062 AAATGCCTTTGCTCTATGGCTGGTCTCTGTGATATGTGAGATCTGGTTTGCCTTATCCTG 1121

Qy       1091 GATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCT 1150
Db       1122 GATTTTGGATCAGTTTCCCAAGTGGTTTCCTGTGAACCGTGAAACCTACCTCGACAGGCT 1181

Qy       1151 GTCACTGAGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGT 1210
Db       1182 TGCTTTAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGCTGCTGTTGACATTTTCGT 1241

Qy       1211 CAGTACGGTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTAT 1270
Db       1242 GAGTACTGTTGACCCCTTGAAGGAGCCACCCCTTGTGACAGCCAACACAGTGCTCTCTAT 1301

Qy       1271 CCTTTCGGTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGC 1330
Db       1302 TCTGGCTGTTGACTACCCAGTTGACAAGGTGTCCTGTTATGTTTCTGATGATGGTGCTGC 1361

Qy       1331 AATGCTAACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTT 1390
Db       1362 TATGTTATCATTTGAATCACTTGCAGAAACATCAGAGTTTGCTCGTAAATGGGTACCATT 1421

Qy       1391 CTGCAAACGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGA 1450
Db       1422 TTGCAAGAAATATAGCATAGAGCCTCGTGCACCAGAATGGTACTTTGCTGCGAAAATAGA 1481

Qy       1451 CTACTTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGA 1510
Db       1482 TTACTTGAAGGATAAAGTTCAGACATCATTTGTCAAAGATCGTAGAGCTATGAAGAGGGA 1541

Qy       1511 GTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGA 1570
Db       1542 ATATGAGGAATTTAAAATCCGAATCAATGCACTTGTTTCCAAAGCCCTAAAATGTCCTGA 1601

Qy       1571 AGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCC 1630
Db       1602 AGAAGGGTGGGTTATGCAAGATGGCACACCGTGGCCTGGAAATAATACAGGGGACCATCC 1661

Qy       1631 TGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACT 1690
Db       1662 AGGAATGATCCAGGTCTTCTTAGGGCAAAATGGTGGACTTGATGCAGAGGGCAATGAGCT 1721

Qy       1691 GCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGC 1750
Db       1722 CCCGCGTTTGGTATATGTTTCTCGAGAAAAGCGACCAGGATTCCAGCACCACAAAAAGGC 1781

Qy       1751 TGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTT 1810
Db       1782 TGGTGCTATGAATGCACTGGTGAGAGTTTCAGCAGTTCTTACCAATGGACCTTTCATCTT 1841

Qy       1811 AAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTT 1870
Db       1842 GAATCTTGATTGTGATCATTACATAAATAACAGCAAAGCCTTAAGAGAAGCAATGTGCTT 1901

Qy       1871 TATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGA 1930
Db       1902 CCTGATGGACCCAAACCTCGGGAAGCAAGTTTGTTATGTTCAGTTCCCACAAAGATTTGA 1961

Qy       1931 TGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGATATCAACAT 1990
Db       1962 TGGTATCGATAAGAACGATAGATATGCTAATCGTAATACCGTGTTCTTTGATATTAACTT 2021

Qy       1991 GAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAG 2050
Db       2022 GAGAGGTTTAGATGGGATTCAAGGACCTGTATATGTCGGAACTGGATGTGTTTTCAACAG 2081

Qy       2051 GCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAA 2110
Db       2082 AACAGCATTATACGGTTATGAACCTCCAATAAAAGTAAA-------------------- 2120

Qy       2111 CTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGAC 2170
Db       2121 ------ACACAAGAAGCCAAGTCTTTTATCTAAGCTCTGTGGTGGATCAAGAAAGAAGAA 2174
```

```
Qy       2171 TACCAAACCCAAAACAGAGAAGAAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATC 2230
Db       2175 TTCCAAAGCTAAGAAAGAGTCGGACAAAAAGAAATCAGGCAGGCATACTGACTCAACTGT 2234

Qy       2231 CCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGC------TGAGAATGA 2284
Db       2235 TCCTGTATTCAACCTCGATGACATAGAAGAGGGAGTTGAAGGTGCTGGTTTTGATGATGA 2294

Qy       2285 AAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTT 2344
Db       2295 AAAGGCGCTCTTAATGTCGCAAATGAGCCTGGAGAAGCGATTTGGACAGTCTGCTGTTTT 2354

Qy       2345 TGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCT 2404
Db       2355 TGTTGCTTCTACCCTAATGGAAAATGGTGGTGTTCCTCCTTCAGCAACTCCAGAAAACTT 2414

Qy       2405 TTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAA 2464
Db       2415 TCTCAAAGAGGCTATCCATGTCATTAGTTGTGGTTATGAGGATAAGTCAGATTGGGGAAT 2474

Qy       2465 AGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCA 2524
Db       2475 GGAGATTGGATGGATCTATGGTTCTGTGACAGAAGATATTCTGACTGGGTTCAAAATGCA 2534

Qy       2525 TTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGC 2584
Db       2535 TGCCCGTGGATGGCGATCCATTTACTGCATGCCTAAGCTTCCAGCTTTCAAGGGTTCTGC 2594

Qy       2585 ACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGA 2644
Db       2595 TCCTATCAATCTTTCAGATCGTCTGAACCAAGTGCTGAGGTGGGCTTTAGGTTCAGTTGA 2654

Qy       2645 GATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAATTTTT 2704
Db       2655 GATTCTCTTCAGTCGGCATTGTCCTATATGGTATGGTTACAATGGGAGGCTAAAATTTCT 2714

Qy       2705 GGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGC 2764
Db       2715 TGAGAGGTTTGCGTATGTGAACACCACCATCTACCCTATCACCTCCATTCCTCTTCTCAT 2774

Qy       2765 TTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAA 2824
Db       2775 GTATTGTACATTGCTAGCCGTTTGTCTCTTCACCAACCAGTTTATTATTCCTCAGATTAG 2834

Qy       2825 TAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCT 2884
Db       2835 TAACATTGCAAGTATATGGTTTCTGTCTCTCTTTCTCTCCATTTTCGCCACGGGTATACT 2894

Qy       2885 AGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGT 2944
Db       2895 AGAAATGAGGTGGAGTGGCGTAGGCATAGACGAATGGTGGAGAAACGAGCAGTTTTGGGT 2954

Qy       2945 CATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGC 3004
Db       2955 CATTGGTGGAGTATCCGCTCATTTATTCGCTGTGTTTCAAGGTATCCTCAAAGTCCTTGC 3014

Qy       3005 TGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGA---GGAGTTCTC 3061
Db       3015 CGGTATTGACACAAACTTCACAGTTACCTCAAAAGCTTCAGATGAAGACGGAGACTTTGC 3074

Qy       3062 AGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTTGCTTCTATT 3121
Db       3075 TGAGCTCTACTTGTTCAAATGGACAACACTTCTGATTCCGCCAACGACGCTGCTCATTGT 3134

Qy       3122 GAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGATATGAGTCATG 3181
Db       3135 AAACTTAGTGGGAGTTGTTGCAGGAGTCTCTTATGCTATCAACAGTGGATACCAATCATG 3194

Qy       3182 GGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTTATCCCTT 3241
Db       3195 GGGACCACTCTTTGGTAAGTTGTTCTTTGCCTTCTGGGTGATTGTTCACTTGTACCCTTT 3254

Qy       3242 TCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCGTCTGGTCCAT 3301
Db       3255 CCTCAAGGGTTTGATGGGTCGACAGAACCGGACTCCTACCATTGTTGTGGTCTGGTCTGT 3314

Qy       3302 TCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCCTTGCGAAGGA 3361
Db       3315 TCTCTTGGCTTCTATCTTCTCGTTGTTGTGGGTTAGGATTGATCCCTTCACTAGCCGAGT 3374

Qy       3362 TGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTG 3397
Db       3375 CACTGGCCCGGACATTCTGGAATGTGGAATCAACTG 3410
```

RESULT 5
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1

Query Match 28.8%; Score 1134.8; DB 3; Length 3328;
Best Local Similarity 66.1%; Pred. No. 5.8e-312;
Matches 1731; Conservative 0; Mismatches 827; Indels 60; Gaps 4;

```
Qy        768 GCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAA 827
Db        488 GCTGAAGACAATGGGAATTCGATTTGGAAGAACAGGGTGGAAAGTTGGAAAGAAAAGAAG 547

Qy        828 GAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTA 887
Db        548 AACAAGAAGAAGAAGCCTGCAACAACTAAGGTTGAAAGAGAGGCTGAAATCCCACCTGAG 607

Qy        888 CCACTAATGGA----------------TGAAGCAAGACAACAACTGTCCAGGAAAATTCCA 932
Db        608 CAACAAATGGAAGATAAACCGGCACCGGATGCTTCCCAGCCCCTCTCGACTATAATTCCA 667

Qy        933 CTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTG 992
Db        668 ATCCCGAAAAGCAGACTTGCACCATACCGAACCGTGATCATTATGCGATTGATCATTCTT 727

Qy        993 GGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTC 1052
Db        728 GGTCTTTTCTTCCATTATCGAGTAACAAACCCCGTTGACAGTGCTTTTGGACTGTGGCTC 787

Qy       1053 ATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAG 1112
Db        788 ACTTCAGTCATATGTGAAATCTGGTTTGCATTTTCCTGGGTGTTGGATCAGTTCCCTAAG 847

Qy       1113 TGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAA 1172
Db        848 TGGTATCCTGTTAACAGGGAAACATACATTGACAGACTATCTGCAAGATATGAAAGAGAA 907

Qy       1173 GGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAG 1232
Db        908 GGTGAACCTGATGAACTTGCTGCAGTTGACTTCTTCGTGAGTACAGTGGATCCATTGAAA 967

Qy       1233 GAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTT 1292
```

```
Db        968 GAGCCTCCATTGATTACTGCCAATACTGTGCTTTCCATCCTTGCCTTGGACTACCCGGTG 1027

Qy       1293 GATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTA 1352

Db       1028 GATAAGGTCTCTTGTTATATATCTGATGATGGTGCGGCCATGCTGACATTTGAATCTCTA 1087

Qy       1353 TCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAA 1412

Db       1088 GTAGAAACAGCCGACTTTGCAAGAAAGTGGGTTCCATTCTGCAAAAAATTTTCCATTGAA 1147

Qy       1413 CCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCA 1472

Db       1148 CCCCGGGCACCTGAGTTTTACTTCTCACAGAAGATTGATTACTTGAAAGATAAAGTGCAG 1207

Qy       1473 GCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGA 1532

Db       1208 CCCTCTTTTGTAAAAGAACGTAGAGCTATGAAAAGAGATTATGAAGAGTACAAAATTCGA 1267

Qy       1533 ATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGAT 1592

Db       1268 ATCAATGCTTTAGTTGCAAAGGCTCAGAAAACACCTGATGAAGGATGGACAATGCAAGAT 1327

Qy       1593 GGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTT 1652

Db       1328 GGAACTTCTTGGCCAGGAAATAACCCGCGTGATCACCCTGGCATGATTCAGGTTTTCCTT 1387

Qy       1653 GGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCT 1712

Db       1388 GGATATAGTGGTGCTCGTGACATCGAAGGAAATGAACTTCCTCGACTGGTTTACGTCTCT 1447

Qy       1713 AGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTC 1772

Db       1448 AGAGAGAAGAGACCTGGCTACCAACACCACAAAAAGGCTGGTGCTGAAAATGCTTTGGTT 1507

Qy       1773 CGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTAC 1832

Db       1508 AGGGTGTCTGCAGTTCTTACAAATGCTCCCTTCATCCTCAATCTTGATTGTGACCACTAT 1567

Qy       1833 ATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGA 1892

Db       1568 GTTAACAATAGCAAGGCAGTTAGGGAGGCAATGTGCTTCTTGATGGACCCACAAGTTGGT 1627

Qy       1893 AAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGA 1952

Db       1628 CGAGATGTATGCTATGTGCAGTTTCCTCAAAGATTTGATGGCATAGATAGGAGTGATCGA 1687

Qy       1953 TATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAG 2012

Db       1688 TATGCCAATAGGAACACAGTTTTCTTTGATGTTAACATGAAAGGTCTTGATGGAATCCAA 1747

Qy       2013 GGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGAT 2072

Db       1748 GGGCCAGTTTATGTGGGAACAGGTTGTGTTTTCAATAGGCAAGCACTTTATGGCTATGGT 1807

Qy       2073 GCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTT 2132

Db       1808 CCACCTTCAATGCCAA----------------------GTTTTCCCAAGTCATCCTCC 1843

Qy       2133 TGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAG 2192

Db       1844 TCATCTTGCTCGTGTTGCTGCCCGGGCAAGAAGGAACCTAAAGATCC------------ 1890

Qy       2193 AAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAA 2252

Db       1891 -----ATCAGAGCTTTATAGGGATGCAAAACGGGAAGAACTTGATGCTGCCATCTTTAAC 1945

Qy       2253 ATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAA 2312

Db       1946 CTTAGGGAAATTGACAATTATGATGAGTATGAAAGATCAATGTTGATCTCTCAAACAAGC 2005

Qy       2313 TTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGT 2372
```

```
Db       2006 TTTGAGAAAACTTTTGGCTTATCTTCAGTCTTCATTGAATCTACACTAATGGAGAATGGA 2065

Qy       2373 GGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGT 2432

Db       2066 GGAGTGGCTGAATCTGCCAACCCTTCCACACTAATCAAGGAAGCAATTCATGTCATCAGC 2125

Qy       2433 TGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTT 2492

Db       2126 TGTGGCTATGAAGAGAAGACTGCATGGGGGAAAGAGATTGGATGGATATATGGTTCAGTC 2185

Qy       2493 ACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGC 2552

Db       2186 ACTGAGGATATCTTAACCGGCTTCAAAATGCACTGCCGAGGATGGAGATCGATTTACTGC 2245

Qy       2553 ATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAC 2612

Db       2246 ATGCCCTTAAGGCCAGCATTCAAAGGATCTGCACCCATCAATCTGTCTGATCGGTTGCAC 2305

Qy       2613 CAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTT 2672

Db       2306 CAGGTTCTTCGATGGGCTCTTGGATCTGTTGAAATTTTCCTAAGCAGGCATTGCCCTCTA 2365

Qy       2673 TGGTATGGGTAT---GGTGGCGGTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCC 2729

Db       2366 TGGTATGGCTTTGGAGGTGGTCGTCTTAAATGGCTTCAAAGACTAGCATATATAAACACC 2425

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789

Db       2426 ATTGTCTATCCTTTCACATCCCTTCCACTCATTGCCTATTGTTCACTACCAGCAATCTGT 2485

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849

Db       2486 CTTCTCACAGGAAAATTTATCATACCAACGCTCTCAAACCTGGCAAGTGTTCTCTTTCTT 2545

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909

Db       2546 GGCCTTTTCCTTTCCATTATCGTGACTGCTGTTCTCGAGCTCCGATGGAGTGGTGTCAGC 2605

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969

Db       2606 ATTGAGGACTTATGGCGTAACGAGCAGTTTTGGGTCATCGGTGGCGTTTCAGCCCATCTC 2665

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029

Db       2666 TTTGCCGTCTTCCAAGGTTTCCTTAAGATGCTTGCGGGCATTGACACCAACTTTACTGTC 2725

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089

Db       2726 ACTGCCAAAGCAGCTGATGATGCAGATTTTGGTGAGCTCTACATTGTGAAATGGACTACA 2785

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149

Db       2786 CTTCTAATCCCTCCAACAACACTCCTCATCGTCAACATGGTTGGTGTCGTTGCCGGATTC 2845

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209

Db       2846 TCCGATGCCCTCAACAAAGGGTACGAAGCTTGGGGACCACTCTTTGGCAAAGTGTTCTTT 2905

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269

Db       2906 TCCTTCTGGGTCATCCTCCATCTTTATCCATTCCTCAAAGGTCTTATGGGACGCCAAAAC 2965

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329

Db       2966 AGGACACCAACCATTGTTGTCCTTTGGTCAGTGTTGTTGGCTTCTGTCTTCTCTCTTGTT 3025

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGG 3367

Db       3026 TGGGTTCGGATCAACCCGTTTGTCAGCACCGCCGATAG 3063
```

RESULT 6
US-09-838-586-1

```
; Sequence 1, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
;  APPLICANT: Stalker, D. et al.
;  TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
;  TITLE OF INVENTION:  Sequences
;  FILE REFERENCE: 15621/02/US
;  CURRENT APPLICATION NUMBER: US/09/838,586
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: 60/029,987
;  PRIOR FILING DATE: 1996-10-29
;  PRIOR APPLICATION NUMBER: 08/960,048
;  PRIOR FILING DATE: 1997-10-29
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3328
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Oligonucleotide
US-09-838-586-1

  Query Match             28.8%;  Score 1134.8;  DB 4;  Length 3328;
  Best Local Similarity   66.1%;  Pred. No. 5.8e-312;
  Matches 1731;  Conservative    0;  Mismatches  827;  Indels   60;  Gaps    4;

Qy        768 GCATATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAA 827
              ||  | |   |||| | |     ||||||| |  || |||| | ||||||  | |    |
Db        488 GCTGAAGACAATGGGAATTCGATTTGGAAGAACAGGGTGGAAAGTTGGAAAGAAAAGAAG 547

Qy        828 GAGAGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGGTGACGATGCTGATCTA 887
                | | || | |  || | |  |    |  | ||     || | |||      |   |
Db        548 AACAAGAAGAAGAAGCCTGCAACAACTAAGGTTGAAAGAGAGGCTGAAATCCCACCTGAG 607

Qy        888 CCACTAATGGA---------------TGAAGCAAGACAACAACTGTCCAGGAAAATTCCA 932
              | || ||||||                || ||    || |  || || |  | ||||||
Db        608 CAACAAATGGAAGATAAACCGGCACCGGATGCTTCCCAGCCCCTCTCGACTATAATTCCA 667

Qy        933 CTTCCATCAAGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTG 992
              | ||   ||||     ||   |||||  | |   | |||||||| ||  |   |  || |
Db        668 ATCCCGAAAAGCAGACTTGCACCATACCGAACCGTGATCATTATGCGATTGATCATTCTT 727

Qy        993 GGGTTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTC 1052
              ||  | ||||||| || |||||  |    | || ||  |   ||| ||||   ||||||||
Db        728 GGTCTTTTCTTCCATTATCGAGTAACAAACCCCGTTGACAGTGCTTTTGGACTGTGGCTC 787

Qy       1053 ATATCTGTTATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAG 1112
              |  || || || |||||||||||||||||  | || ||| |  | |||||  ||||| |||
Db        788 ACTTCAGTCATATGTGAAATCTGGTTTGCATTTTCCTGGGTGTTGGATCAGTTCCCTAAG 847

Qy       1113 TGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAA 1172
              ||||  ||| || | || || ||  | ||| | || ||    || |  || |   |||
Db        848 TGGTATCCTGTTAACAGGGAAACATACATTGACAGACTATCTGCAAGATATGAAAGAGAA 907

Qy       1173 GGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAG 1232
              ||  | ||   | |||||||| || |||| ||||| || ||||| || ||||| || ||
Db        908 GGTGAACCTGATGAACTTGCTGCAGTTGACTTCTTCGTGAGTACAGTGGATCCATTGAAA 967

Qy       1233 GAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTT 1292
              || ||||| ||| | ||  | |||||||| || || ||||||  |  |||| || || ||
Db        968 GAGCCTCCATTGATTACTGCCAATACTGTGCTTTCCATCCTTGCCTTGGACTACCCGGTG 1027

Qy       1293 GATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTA 1352
              |||||||| ||||| ||| | |||||||||||||| || ||||| || ||||||  |  ||
Db       1028 GATAAGGTCTCTTGTTATATATCTGATGATGGTGCGGCCATGCTGACATTTGAATCTCTA 1087

Qy       1353 TCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAA 1412
                ||||||  | || ||||||||  || ||||||||  |||||||||      |||||||
Db       1088 GTAGAAACAGCCGACTTTGCAAGAAAGTGGGTTCCATTCTGCAAAAAATTTTCCATTGAA 1147

Qy       1413 CCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCA 1472
              || || ||  || ||||  ||||||  ||||||||| || ||||||||||  || |||
```

```
Db       1148 CCCCGGGCACCTGAGTTTTACTTCTCACAGAAGATTGATTACTTGAAAGATAAAGTGCAG 1207

Qy       1473 GCAAACTTTGTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGA 1532
                |    ||||| |  ||  | ||||| ||||| ||||| ||||| || | |||  |  ||
Db       1208 CCCTCTTTTGTAAAAGAACGTAGAGCTATGAAAAGAGATTATGAAGAGTACAAAATTCGA 1267

Qy       1533 ATCAATGCCTTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGAT 1592
              |||||||| |||||||| || || ||||||    ||||| |||||||||||||||||||||
Db       1268 ATCAATGCTTTAGTTGCAAAGGCTCAGAAAACACCTGATGAAGGATGGACAATGCAAGAT 1327

Qy       1593 GGAACCCCCTGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTT 1652
              |||||  | ||||| ||||| ||    |||||||| ||||| |||||||||| ||||||
Db       1328 GGAACTTCTTGGCCAGGAAATAACCCGCGTGATCACCCTGGCATGATTCAGGTTTTCCTT 1387

Qy       1653 GGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCT 1712
              ||  | || || |  | ||||   || |||||||||| || ||| ||||||| || |||
Db       1388 GGATATAGTGGTGCTCGTGACATCGAAGGAAATGAACTTCCTCGACTGGTTTACGTCTCT 1447

Qy       1713 AGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTC 1772
              |||||||| |||| ||||| | || || || || |||||||||    ||||| |||||
Db       1448 AGAGAGAAGAGACCTGGCTACCAACACCACAAAAAGGCTGGTGCTGAAAATGCTTTGGTT 1507

Qy       1773 CGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTAC 1832
                | || |||||  || || ||||||||||  |   |  | ||  | |||||||| |||||
Db       1508 AGGGTGTCTGCAGTTCTTACAAATGCTCCCTTCATCCTCAATCTTGATTGTGACCACTAT 1567

Qy       1833 ATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGA 1892
                | |||||  ||||||||  | | |||  |||||||| ||  |||||||||   | | ||
Db       1568 GTTAACAATAGCAAGGCAGTTAGGGAGGCAATGTGCTTCTTGATGGACCCACAAGTTGGT 1627

Qy       1893 AAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGA 1952
                | || |||||||| |||||  |||||||||||||||||| || |||  |   |||  |||
Db       1628 CGAGATGTATGCTATGTGCAGTTTCCTCAAAGATTTGATGGCATAGATAGGAGTGATCGA 1687

Qy       1953 TATGCTAACCGGAATGTTGTCTTTTTTGATATCAACATGAAAGGTTTGGATGGTATTCAG 2012
              |||||  ||  ||||    || || ||||||  | ||||||||||||  | |||||  || ||
Db       1688 TATGCCAATAGGAACACAGTTTTCTTTGATGTTAACATGAAAGGTCTTGATGGAATCCAA 1747

Qy       2013 GGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTATATGGTTATGAT 2072
              || ||| ||||||| || || || ||||| || |  |||||  ||| | |||||  |||| |
Db       1748 GGGCCAGTTTATGTGGGAACAGGTTGTGTTTTCAATAGGCAAGCACTTTATGGCTATGGT 1807

Qy       2073 GCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTT 2132
                | ||   ||    |                          | |  |||||||  | ||
Db       1808 CCACCTTCAATGCCAA-----------------------GTTTTCCCAAGTCATCCTCC 1843

Qy       2133 TGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAG 2192
              |  | |||||  | | ||     |    | ||| |  |||    | ||
Db       1844 TCATCTTGCTCGTGTTGCTGCCCGGGCAAGAAGGAACCTAAAGATCC------------- 1890

Qy       2193 AAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAA 2252
                  | |    |||  | | | | | |      ||    | ||    |||  |   | ||
Db       1891 -----ATCAGAGCTTTATAGGGATGCAAAACGGGAAGAACTTGATGCTGCCATCTTTAAC 1945

Qy       2253 ATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAA 2312
                ||   |||  ||         | |||| ||||||   |      |  |   ||||  ||
Db       1946 CTTAGGGAAATTGACAATTATGATGAGTATGAAAGATCAATGTTGATCTCTCAAACAAGC 2005

Qy       2313 TTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGT 2372
              || || || |  ||||||  |||||| || ||  ||  ||| |||||  | ||||||||
Db       2006 TTTGAGAAAACTTTTGGCTTATCTTCAGTCTTCATTGAATCTACACTAATGGAGAATGGA 2065

Qy       2373 GGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGT 2432
              |||      |  ||| |  ||| |  | ||  | || |||||  || ||||||||  ||
Db       2066 GGAGTGGCTGAATCTGCCAACCCTTCCACACTAATCAAGGAAGCAATTCATGTCATCAGC 2125

Qy       2433 TGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCTATGGATCAGTT 2492
              ||||| |||||||| ||||| |  ||||| |||||||||||  ||||| ||||| |||||
Db       2126 TGTGGCTATGAAGAGAAGACTGCATGGGGGAAAGAGATTGGATGGATATATGGTTCAGTC 2185

Qy       2493 ACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGC 2552
              || || |||||  |||| || ||||| |||||  ||  |  || ||| | || ||||||||||
Db       2186 ACTGAGGATATCTTAACCGGCTTCAAAATGCACTGCCGAGGATGGAGATCGATTTACTGC 2245
```

```
Qy       2553 ATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAC 2612
Db       2246 ATGCCCTTAAGGCCAGCATTCAAAGGATCTGCACCCATCAATCTGTCTGATCGGTTGCAC 2305

Qy       2613 CAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTT 2672
Db       2306 CAGGTTCTTCGATGGGCTCTTGGATCTGTTGAAATTTTCCTAAGCAGGCATTGCCCTCTA 2365

Qy       2673 TGGTATGGGTAT---GGTGGCGGTCTGAAATTTTTGGAAAGATTTCCTACATCAACTCC 2729
Db       2366 TGGTATGGCTTTGGAGGTGGTCGTCTTAAATGGCTTCAAAGACTAGCATATATAAACACC 2425

Qy       2730 ATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGT 2789
Db       2426 ATTGTCTATCCTTTCACATCCCTTCCACTCATTGCCTATTGTTCACTACCAGCAATCTGT 2485

Qy       2790 TTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATG 2849
Db       2486 CTTCTCACAGGAAAATTTATCATACCAACGCTCTCAAACCTGGCAAGTGTTCTCTTTCTT 2545

Qy       2850 TCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGA 2909
Db       2546 GGCCTTTTCCTTTCCATTATCGTGACTGCTGTTCTCGAGCTCCGATGGAGTGGTGTCAGC 2605

Qy       2910 ATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTC 2969
Db       2606 ATTGAGGACTTATGGCGTAACGAGCAGTTTGGGTCATCGGTGGCGTTTCAGCCCATCTC 2665

Qy       2970 TTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTG 3029
Db       2666 TTTGCCGTCTTCCAAGGTTTCCTTAAGATGCTTGCGGGCATTGACACCAACTTTACTGTC 2725

Qy       3030 ACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
Db       2726 ACTGCCAAAGCAGCTGATGATGCAGATTTTGGTGAGCTCTACATTGTGAAATGGACTACA 2785

Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
Db       2786 CTTCTAATCCCTCCAACAACACTCCTCATCGTCAACATGGTTGGTGTCGTTGCCGGATTC 2845

Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
Db       2846 TCCGATGCCCTCAACAAAGGGTACGAAGCTTGGGGACCACTCTTTGGCAAAGTGTTCTTT 2905

Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
Db       2906 TCCTTCTGGGTCATCCTCCATCTTTATCCATTCCTCAAAGGTCTTATGGACGCCAAAAC 2965

Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
Db       2966 AGGACACCAACCATTGTTGTCCTTTGGTCAGTGTTGTTGGCTTCTGTCTTCTCTCTTGTT 3025

Qy       3330 TGGGTTCGGATTGATCCTTTCCTTGCGAAGGATGATGG 3367
Db       3026 TGGGTTCGGATCAACCCGTTTGTCAGCACCGCCGATAG 3063
```

RESULT 7
US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2248
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1887)
US-09-221-013A-1

```
  Query Match             22.5%;  Score 884.6;  DB 4;  Length 2248;
  Best Local Similarity   68.5%;  Pred. No. 6.2e-241;
  Matches 1277;  Conservative    0;  Mismatches  559;  Indels   27;  Gaps    3;

Qy       1495 GAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAG 1554
Db          2 GAGCTATGAAGAGAGAGTATGAAGAGTTTAAAGTGAGGATAAATGCTCTTGTTGCCAAAG 61

Qy       1555 CCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACA 1614
Db         62 CACAGAAAATCCCTGGAGAAGGCTGGACAATGCAGGATGGTACTCCCTGGCCTGGTAACA 121

Qy       1615 ATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACT 1674
Db        122 ACACTAGAGATCATCCTGGAATGATACAGGTGTTCTTAGGCCATAGTGGGGGTCTGGATA 181

Qy       1675 GTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATA 1734
Db        182 CCGATGGAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAAAGCGGCCTGGATTTC 241

Qy       1735 ACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAA 1794
Db        242 AACACCACAAAAAGGCTGGAGCTATGAATGCATCGATCCGTGTATCTGCTGTTCTTACCA 301

Qy       1795 ATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAA 1854
Db        302 ATGGAGCATATCTTTTGAACGTGGATTGTGATCATTACTTTAATAACAGTAAGGCTATTA 361

Qy       1855 AGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGT 1914
Db        362 AAGAAGCTATGTGTTTCATGATGGACCCGGCTATTGGAAAGAAGTGCTGCTATGTCCAGT 421

Qy       1915 TCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCT 1974
Db        422 TCCCTCAACGTTTTGACGGTATTGATTTGCACGATCGATATGCCAACAGGAATATAGTCT 481

Qy       1975 TTTTGATATCAACATGAAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTG 2034
Db        482 TTTTCGATATTAACATGAAGGGGTTGGATGGTATCCACGGTCCAGTATATGTGGGTACTG 541

Qy       2035 GATGTGTATTTAGAAGGCAGGCATTATATGGTTATGATGCCCCCAAAACAAAGAAGCCAC 2094
Db        542 GTTGTTGTTTTAATAGGCAGGCTCTATATGGGTATGATCCTGTTTTGACGGAAGAAGATT 601

Qy       2095 CATCAAGGACTTGCAACTGCTGGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATA 2154
Db        602 TAGAACCAAATATTA--------------------TTGTCAAGAGCTGTTGCGGGTCAA 640

Qy       2155 GGAAGCAAAAGAAGACTACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGA 2214
Db        641 GGAAGAAAGGTAAAAGTAGCAA---GAAGTATAACTACGAAAAGAGGAGAGGCATCAACA 697

Qy       2215 AAGAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAG 2274
Db        698 GAAGTGACTCCAATGCTCCACTTTTCAATATGGAGGACATCGATGAGGGTTTGAAGGTT 757

Qy       2275 CTGAGAATGAAAAGGCCGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAAT 2334
Db        758 ATGATGATGAGAGGTCTATTCTAATGTCCCAGAGGAGTGTAGAGAAGCGTTTTGGTCAGT 817
```

```
Qy       2335 CTTCTGTTTTTGTTACATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTC 2394
Db        818 CGCCGGTATTTATTGCGGCAACCTTCATGGAACAAGGCGGCATTCCACCAACAACCAATC 877

Qy       2395 CTGCTTCTCTTTTGAAAGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAG 2454
Db        878 CCGCTACTCTTCTGAAGGAGGCTATTCATGTTATAAGCTGTGGTTACGAAGACAAGACTG 937

Qy       2455 ACTGGGGAAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTT 2514
Db        938 AATGGGGCAAAGAGATTGGTTGGATCTATGGTTCCGTGACGGAAGATATTCTTACTGGGT 997

Qy       2515 TCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCA 2574
Db        998 TCAAGATGCATGCCCGGGGTTGGATATCGATCTACTGCAATCCTCCACGCCCTGCGTTCA 1057

Qy       2575 AAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTG 2634
Db        1058 AGGGATCTGCACCAATCAATCTTTCTGATCGTTTGAACCAAGTTCTTCGATGGGCTTTGG 1117

Qy       2635 GGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTC 2694
Db        1118 GATCTATCGAGATTCTTCTTAGCAGACATTGTCCTATCTGGTATGGTTACCATGGAAGGT 1177

Qy       2695 TGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTC 2754
Db        1178 TGAGACTTTTGGAGAGGATCGCTTATATCAACACCATCGTCTATCCTATTACATCCATCC 1237

Qy       2755 CCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTC 2814
Db        1238 CTCTTATTGCGTATTGTATTCTTCCCGCTTTTTGTCTCATCACCGACAGATTCATCATAC 1297

Qy       2815 CAGAGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTA 2874
Db        1298 CCGAGATAAGCAACTACGCGAGTATTTGGTTCATTCTACTCTTCATCTCAATTGCTGTGA 1357

Qy       2875 CGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGC 2934
Db        1358 CTGGAATCCTGAAACTGAAATGGAACGGTGTGAGCATTGAGGATTGGTGGAGGAACAACC 1417

Qy       2935 AGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCA 2994
Db        1418 AGTTCTGGGTCATTGGTGGCACATCCACCCATCTTTTTGCTGTCTTCCAAGGTCTACTTA 1477

Qy       2995 AGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGA---GATGATG 3051
Db        1478 AGGTTCTTGCTGGTATCAACACCAACTTCACCGTTACATCTAAAGCCACAAACAAAAATG 1537

Qy       3052 AGGAGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCT 3111
Db        1538 GGGATTTTGCAAAACTCTACATCTTCAAATGGACAGCTCTTCTCATTCCACCAACCACCG 1597

Qy       3112 TGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTTTCAAATGCGATCAATAACGGAT 3171
Db        1598 TCCTACTTGTGAACCTCATAGGCATTGTGGCTGGTGTCTCTTATGCTGTAAACAGTGGCT 1657

Qy       3172 ATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTGATTGTCCATC 3231
Db        1658 ACCAGTCGTGGGGTCCGCTTTTCGGGAAGCTCTTCTTCGCCTTATGGGTTATTGCCCATC 1717

Qy       3232 TTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCATCG 3291
Db        1718 TCTACCCTTTCTTGAAAGGTCTGTTGGGAAGACAAAACCGAACACCAACCATCGTCATTG 1777

Qy       3292 TCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTCGGATTGATCCTTTCC 3351
Db        1778 TCTGGTCTGTTCTTCTCGCCTCCATCTTCTCGTTGCTTTGGGTCAGGATCAATCCCTTTG 1837

Qy       3352 TTG 3354
Db        1838 TGG 1840
```

```
RESULT 8
US-09-221-013A-13
; Sequence 13, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 1741
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(1741)
US-09-221-013A-13

  Query Match            16.9%;  Score 664.2;  DB 4;  Length 1741;
  Best Local Similarity   64.5%;  Pred. No. 2.5e-178;
  Matches 1087;  Conservative    0;  Mismatches  563;  Indels   35;  Gaps    5;

Qy        117 GGCGGCGGAAGTGGAGGGGAGGAAGCGATGGAGGCGAGCGCCGGGCTGGTGGCCGGCTCC 176
Db         76 GGAGGAGGAGGAGGAAGGGAGGA--GGATGGCGGCGAACGCGGGGATGGTGGCGGGATCC 133

Qy        177 CACAACCGCAACGAGCTCGTCGTCATCCGCCGCGACGGCGA------TCCCGGGCCGAAG 230
Db        134 CGCAACCGGAACGAGTTCGTCATGATCCGCCCCGACGGCGACGCGCCACCGCCGGCTAAG 193

Qy        231 CCGCCGCGGGAGCAGAACGGGCAGGTGTGCCAGATTTGCGGCGACGACGTCGGCCTTGCC 290
Db        194 CCAGGGAAGAGTGTGAATGGTCAGGTCTGCCAGATTTGTGGCGACACTGTTGGCGTCTCG 253

Qy        291 CCCGGCGGGGACCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCCGTCTGCCGGGACTGC 350
Db        254 GCCACCGGCGACGTCTTTGTTGCCTGCAATGAGTGCGCCTTCCCGGTCTGCCGCCCTTGC 313

Qy        351 TACGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAAGACTCGATACAAG 410
Db        314 TACGAGTACGAACGCAAGGAAGGGAACCAGTGCTGCCCCCAGTGCAAGACTAGATACAAG 373

Qy        411 CGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACCTG 470
Db        374 AGGCACAAAGGTTGCCCTAGAGTTCAGGGCGATGAGGAAGAAGAAGATGTTGATGACCTG 433

Qy        471 GACAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTAC 530
Db        434 GACAATGAATTCCATTATAA----------------GCATGGCAATGGCAAAGGTCCAG 476

Qy        531 GGCCACATGAGCTACGGCCGTGGAGGTGACCCTAATGGCGCGCCACAAGCTTTCCAGCTC 590
Db        477 AGTGGCA-GATACAGAGACAGGGGGAAGATGTTGACCTGTCTTCATCTTCTCGCCACGAA 535

Qy        591 AACCCCAATGTTCCACTCCTCACCAACGGGCAAATGGTGGATGACATCCCACCGGAGCAG 650
Db        536 CAACATCGGATTCCCCGTCTGACAAGTGGGCAACAGATCTCAGGAGAGATCCCTGATGCT 595

Qy        651 CACGCGCTGGTGCCTTCTTTCATGGGTGGTGGGGGAAAGAGGATACATCCCCTTCCTTAT 710
Db        596 TCCCCCGATCGCCATTCTATCCGCAGCGGAACATCAAGCTATGTTGATCCAAGTGTTCCA 655

Qy        711 GCGGATCCCAGCTTACCTGTGCAACCCAGGTCTATGGACCCATCCAAGGATCTTGCTGCA 770
```

```
Db       656 GTTCCTGTGAGGATTGTGGACCCCTCCAAGGACTTGAATTCCTATGGGATTAACAGTGTT 715

Qy       771 TATGGGTATGGTAGTGTTGCTTGGAAGGAACGGATGGAGAATTGGAAGCAGAGACAAGAG 830

Db       716 GACTGGCAAGAAAGAGTTGCCAGCTGGAGGAACAAGCAGGACAAAAATATGATGCAGGTA 775

Qy       831 AGGATGCACCAGACGGGGAATGATGGTGGTGGTGATGATGG---------TGACGATGCT 881

Db       776 GCTAATAAATATCCAGAGGCAAGAGGGGGAGACATGGAAGGGACTGGTTCAAATGGTGAA 835

Qy       882 GATCTACCACTAATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCA 941

Db       836 GATATCCAAATGGTTGATGATGCACGTCTACCTCTGAGCCGCATAGTGCCTATCCCTTCA 895

Qy       942 AGCCAGATTAATCCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTC 1001

Db       896 AACCAGCTCAACCTTTACCGGATTGTTATCATTCTCCGTCTTATCATCCTGATGTTCTTC 955

Qy      1002 TTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTT 1061

Db       956 TTCCAATATCGTGTCACTCATCCAGTGCGGGATGCTTATGGATTGTGGCTAGTATCTGTT 1015

Qy      1062 ATCTGTGAAATCTGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCT 1121

Db      1016 ATCTGTGAAATTTGGTTGCCCTTATCCTGGCTCCTAGATCAATTCCCAAAGTGGTACCCG 1075

Qy      1122 ATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTCGACAAGGAAGGCCAGCCA 1181

Db      1076 ATAAACCGTGAAACATACCTTGACAGGCTTGCATTGAGATATGATAGGGAGGGAGAGCCA 1135

Qy      1182 TCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAACCTCCT 1241

Db      1136 TCACAGCTTGCTCCCATTGATGTCTTTGTCAGTACGGTGGATCCACTAAAGGAACCTCCT 1195

Qy      1242 TTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTATCCTGTTGATAAGGTT 1301

Db      1196 CTGATCACAGCAAACACTGTTTTGTCCATTCTGGCTGTGGATTACCCTGTTGACAAAGTG 1255

Qy      1302 TCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGAAGCATTATCTGAAACA 1361

Db      1256 TCATGCTATGTTTCTGACGATGGTTCAGCTATGTTAACTTTTGAGGCTCTGTCAGAAACT 1315

Qy      1362 TCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAATATTGAACCTCGCGCT 1421

Db      1316 GCAGAATTTGCTAGGAAGTGGGTTCCGTTTTGCAAGAAGCACAATATTGAACCACGAGCT 1375

Qy      1422 CCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAAGGTGGCAGCAAACTTT 1481

Db      1376 CCAGAGTTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAAAATCCAACCTTCCTTT 1435

Qy      1482 GTTAGGGAGAGGAGAGCAATGAAGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCC 1541

Db      1436 GTTAAAGAAAGGCGGGCAATGAAGAGAGAGTATGAAGAATTCAAGGTACGGATCAATGCT 1495

Qy      1542 TTAGTTGCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCC 1601

Db      1496 CTTGTTGCGAAGGCACAAAAAGTACCTGAAGAGGGGTGGACCATGGCTGATGGCACTGCT 1555

Qy      1602 TGGCCTGGAAACAATGTTCGTGATCATCCTGGAATGATTCAGGTCTTCCTTGGCCAAAGC 1661

Db      1556 TGGCCTGGGAATAACCCAAGGGATCACCCTGGCATGATTCAGGTGTTCTTGGGGCACAGT 1615

Qy      1662 GGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAA 1721

Db      1616 GGTGGGCTTGACACTGATGGTAACGAGTTGCCACGGCTTGTCTACGTCTCTCGTGAAAAG 1675

Qy      1722 CGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGGTCCGAGTCTCT 1781

Db      1676 AGGCCAGGATTCCAGCATCACAAGAAGGCTGGTGCAATGAATGCATTGATTCGTGTATCT 1735

Qy      1782 GCTGT 1786
```

```
Db       1736 GCTGT 1740


RESULT 9
US-09-221-013A-3
; Sequence 3, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 8411
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-221-013A-3

  Query Match             10.1%;  Score 397.8;  DB 4;  Length 8411;
  Best Local Similarity   54.1%;  Pred. No. 4.2e-102;
  Matches 1540;  Conservative    0;  Mismatches  657;  Indels  648;  Gaps   10;

Qy       1158 AGGTTCGACAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACG 1217

Db       4781 AGATATGATCGAGACGGTGAACCATCACAGCTCGTTCCTGTTGATGTGTTTGTTAGTACA 4840

Qy       1218 GTTGATCCCTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCG 1277

Db       4841 GTGGACCCATTGAAAGAGCCTCCCCTTGTTACAGCAAACACAGTTCTCTCGATTCTTTCT 4900

Qy       1278 GTGGATTATCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTA 1337

Db       4901 GTGGACTACCCGGTAGATAAAGTAGCCTGTTATGTTTCAGATGATGGTTCAGCTATGCTT 4960

Qy       1338 ACGTTTGAAGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAA 1397

Db       4961 ACCTTTGAATCCCTTTCTGAAACCGCTGAGTTTGCAAAGAAATGGGTACCATTTTGCAAG 5020

Qy       1398 CGGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTG 1457

Db       5021 AAATTCAACATTGAACCTAGGGCCCCTGAATTCTATTTTGCCCAGAAGATAGATTACTTG 5080

Qy       1458 AAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGA-------------- 1503

Db       5081 AAGGACAAGATCCAACCGTCTTTTGTTAAAGAGCGACGAGCTATGAAGGTCATTTGAAAA 5140

Qy       1504 ---------------------------------------------------------- 1503

Db       5141 GTCCACCTGCTTCTCATCCATACGGCAAAGAGATTGACTGACTTTTTCTTTGGTTTGTAT 5200

Qy       1504 ----AGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCCAG 1559

Db       5201 TGACAGAGAGAGTATGAAGAGTTTAAAGTGAGGATAAATGCTCTTGTTGCCAAAGCACAG 5260

Qy       1560 AAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCTGGAAACAATGTT 1619

Db       5261 AAAATCCCTGAAGAAGGCTGGACAATGCAGGATGGTACTCCCTGGCCTGGTAACAACACT 5320

Qy       1620 CGTGATCATCCTGGAATGA----------------------------------------- 1638

Db       5321 AGAGATCATCCTGGAATGATACAGGTACAGTGTGGCAATCCCTTGATTGTGACAGAGAGG 5380

Qy       1639 -----------------------------------TTCAGGTCTTCCTTGGCCAA 1658
```

```
Db       5381 ATAACGTAAAGGAAACATGTTTACATCGTTTTGTTTCAATTTCAGGTGTTCTTAGGCCAT 5440
Qy       1659 AGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAG 1718
Db       5441 AGTGGGGGTCTGGATACCGATGGAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAA 5500
Qy       1719 AAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTATGAATGCATTGG-------- 1770
Db       5501 AAGCGGCCTGGATTTCAACACCACAAAAAGGCTGGAGCTATGAATGCATTGGTTTGTTAA 5560
Qy       1771 ------------------------------------------------------------ 1770
Db       5561 CTTTCAGAATCCTATTGTGTCCTCTATTTTATTCTCTTGTTCACTGCCTAAGAAACGTTC 5620
Qy       1771 ------------------------------------------------------------ 1770
Db       5621 TTCTTGTGTAGCCGTTGCTTCACATTCTTTTTTTTCTAGGCTATGTGTTCTCTCCTAATT 5680
Qy       1771 ----------------------TCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTT 1807
Db       5681 TAGTATCTCTTTACTTTGACAGATCCGTGTATCTGCTGTTCTTACCAATGGAGCATATCT 5740
Qy       1808 GTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTG 1867
Db       5741 TTTGAACGTGGATTGTGATCATTACTTTAATAACAGTAAGGCTATTAAAGAAGCTATGTG 5800
Qy       1868 TTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATGTACAGTTCCCTCAAAGATT 1927
Db       5801 TTTCATGATGGACCCGGCTATTGGAAAGAAGTGCTGCTATGTCCAGTTCCCTCAACGTTT 5860
Qy       1928 TGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGAT----- 1982
Db       5861 TGACGGTATTGATTTGCACGATCGATATGCCAACAGGAATATAGTCTTTTTCGATGTGAG 5920
Qy       1983 ------------------------------------------------------------ 1982
Db       5921 TATCACTTCCCCATTGTCTTTTGTTTCTCTTTTGTTCATATTTTGGTTGGATTTACTCGT 5980
Qy       1983 ----------------------------------------ATCAACATGAAAGGTTTGG 2001
Db       5981 TTCTGCTATGGCCTGACTTGGATATTTGTTCTCTTGGGCAGATTAACATGAAGGGGTTGG 6040
Qy       2002 ATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCATTAT 2061
Db       6041 ATGGTATCCAGGGTCCAGTATATGTGGGTACTGGTTGTTGTTTTAATAGGCAGGCTCTAT 6100
Qy       2062 ATGGTTATGATGC------CCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCT 2115
Db       6101 ATGGGTATGATCCTGTTTTGACGGAAGAAGATTTAGAACCAAATATTATTGTCAAGAGCT 6160
Qy       2116 GGCCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCA 2175
Db       6161 GTTGCGGGTCAAGGAAGAAAGGTAAAAGTAGCAAGAAGTATAACTACGAAAAGAGGAGAG 6220
Qy       2176 AACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCA-------- 2227
Db       6221 GCATCAACAGAAGTGACTCCAATGCTCCACTTTTCAATATGGAGGACATCGATGAGGGTT 6280
Qy       2228 ----------------------------------------ATCCCCTGCATATGCT 2243
Db       6281 TTGAAGGTTTGATTGAGCTGATTGTGTAATAACATCACTTCTTTATGTAATGATTTATGT 6340
Qy       2244 CTTGGTGAAATTGACGAA-----GCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTG 2298
Db       6341 GATGGTGAAATCTTACAATCCTTGTTTATGCAGGTTATGATGATGAGAGGTCTATTCTAA 6400
Qy       2299 TAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATCCACAC 2358
Db       6401 TGTCCCAGAGGAGTGTAGAGAAGCGTTTTGGTCAGTCGCCGGTATTTATTGCGGCAACCT 6460
Qy       2359 TTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGAAGCTA 2418
```

```
Db       6461 TCATGGAACAAGGCGGCATTCCACCAACAACCAATCCCGCTACTCTTCTGAAGGAGGCTA 6520
Qy       2419 TACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAG----------- 2466
Db       6521 TTCATGTTATAAGCTGTGGTTACGAAGACAAGACTGAATGGGGCAAAGAGGTCAGTTTTC 6580
Qy       2467 ------------------------------------------------------------ 2466
Db       6581 AAATGCAGCTACAGAATCTTCTTATGTTCTCTTTCTTACCTGTTTGATGACATCTTATTT 6640
Qy       2467 ------------AGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTGGTT 2514
Db       6641 GGCACTTTTGTTAGATTGGTTGGATCTATGGTTCCGTGACGGAAGATATTCTTACTGGGT 6700
Qy       2515 TCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCA 2574
Db       6701 TCAAGATGCATGCCCGGGGTTGGATATCGATCTACTGCAATCCTCCACGCCCTGCGTTCA 6760
Qy       2575 AAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTG 2634
Db       6761 AGGGATCTGCACCAATCAATCTTTCTGATCGTTTGAACCAAGTTCTTCGATGGGCTTTGG 6820
Qy       2635 GGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCGGTC 2694
Db       6821 GATCTATCGAGATTCTTCTTAGCAGACATTGTCCTATCTGGTATGGTTACCATGGAAGGT 6880
Qy       2695 TGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTATTC 2754
Db       6881 TGAGACTTTTGGAGAGGATCGCTTATATCAACACCATCGTCTATCCTATTACATCCATCC 6940
Qy       2755 CCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCACTC 2814
Db       6941 CTCTTATTGCGTATTGTATTCTTCCCGCTTTTTGTCTCATCACCGACAGATTCATCATAC 7000
Qy       2815 CAGAGCTGAATAATGTT---------------------------------------- 2831
Db       7001 CCGAGGTTTGTAAAACTGACCACACTGCTATTTACTATTTGAATCCCATTTTGTGAATGC 7060
Qy       2832 -----------------------------------------GCCAGCCTGTGGTTCATGTCA 2852
Db       7061 ATTTTTTTGTCATCATCATTGTTGCAGATAAGCAACTACGCGAGTATTTGGTTCATTCTA 7120
Qy       2853 CTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGGAATT 2912
Db       7121 CTCTTCATCTCAATTGCTGTGACTGGAATCCTGGAGCTGAGATGGAGCGGTGTGAGCATT 7180
Qy       2913 GATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTT 2972
Db       7181 GAGGATTGGTGGAGGAACGAGCAGTTCTGGGTCATTGGTGGCACATCCGCCCATCTTTTT 7240
Qy       2973 GCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGTGACA 3032
Db       7241 GCTGTCTTCCAAGGTCTACTTAAGGTTCTTGCTGGTATCGACACCAACTTCACCGTTACA 7300
Qy       3033 TCAAAGG---GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTACC 3089
Db       7301 TCTAAAGCCACAGACGAAGATGGGGATTTTGCAGAACTCTACATCTTCAAATGGACAGCT 7360
Qy       3090 TTATTGATACCTCCTACCACCTTGCTTCTATTGAACTTCATTGGTGTGGTCGCTGGCGTT 3149
Db       7361 CTTCTCATTCCACCAACCACCGTCCTACTTGTGAACCTCATAGGCATTGTGGCTGGTGTC 7420
Qy       3150 TCAAATGCGATCAATAACGGATATGAGTCATGGGGCCCCCTCTTTGGGAAGCTATTCTTT 3209
Db       7421 TCTTATGCTGTAAACAGTGGCTACCAGTCGTGGGGTCCGCTTTTCGGGAAGCTCTTCTTC 7480
Qy       3210 GCATTTTGGGTGATTGTCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAC 3269
Db       7481 GCCTTATGGGTTATTGCCCATCTCTACCCTTTCTTGAAAGGTCTGTTGGGAAGACAAAAC 7540
Qy       3270 AGGACACCAACGATTGTCATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTT 3329
Db       7541 CGAACACCAACCATCGTCATTGTCTGGTCTGTTCTTCTCGCCTCCATCTTCTCGTTGCTT 7600
```

```
Qy       3330 TGGGTTCGGATTGATCCTTTCCTTG 3354
             |||||  ||||  |||| ||  | |
Db       7601 TGGGTCAGGATCAATCCCTTTGTGG 7625
```

RESULT 10
US-09-221-013A-4
; Sequence 4, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(5009)
; OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4

Query Match 8.0%; Score 316.8; DB 4; Length 5009;
Best Local Similarity 53.5%; Pred. No. 3.2e-79;
Matches 1384; Conservative 0; Mismatches 542; Indels 662; Gaps 8;

```
Qy        894 ATGGATGAAGCAAGACAACAACTGTCCAGGAAAATTCCACTTCCATCAAGCCAGATTAAT 953
              |||||||||| ||||||||  || || || ||  | ||  |||  ||||||   || |||
Db       2422 ATGGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACCTATTCGTTCAAGCAGAATAAAT 2481

Qy        954 CCATATAGGATGATTATCATTATTCGGCTTGTGGTTTTGGGGTTCTTCTTCCACTACCGA 1013
              || || ||||||  | ||  |   ||| || | | || | ||  | ||||| || ||  ||
Db       2482 CCTTACAGGATGTTAATTCTGTGTCGCCTCGCGATTCTTGGTCTTTTCTTTCATTATAGA 2541

Qy       1014 GTGATGCATCCGGTGAATGATGCATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAATC 1073
                |  | ||||| || |||||||||| ||  || ||| | |  || ||||| || || ||
Db       2542 ATTCTCCATCCAGTCAATGATGCATATGGATTATGGTTAACGTCAGTTATATGCGAGATA 2601

Qy       1074 TGGTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGGTTCCCTATTGAGAGAGAG 1133
              |||||||| |||||||||||||||||||||||| || ||||  |||||| ||  | ||
Db       2602 TGGTTTGCAGTGTCTTGGATTCTTGATCAATTCCCCAAATGGTATCCTATAGAACGTGAA 2661

Qy       1134 ACTTACCTAGACCGGCTGTCAC---------------------------------------- 1155
              || ||||| ||  | || || |
Db       2662 ACATACCTCGATAGACTCTCTCTCAGGTAACATAAACCCTGAAAAGTTCTTGTCTGCAAA 2721

Qy       1156 ----------------------------------------------------TGAGGTTCGA 1165
                                                                   | |||| |||
Db       2722 TATTCATTTTTTACATTCCCAAAAATTTTTGAAACTCTATTTTTCTTACATAAGGTACGA 2781

Qy       1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225
              ||||||||  | || ||   | | || ||  ||||| | ||||| ||||| || |||||
Db       2782 GAAGGAAGGAAAACCGTCAGGATTAGCACCTGTTGATGTTTTTGTTAGTACAGTGGATCC 2841

Qy       1226 CTTAAAGGAACCTCCTTTGGTCACAACAAATACTGTTCTATCTATCCTTTCGGTGGATTA 1285
              || || || || || ||| | ||| |||| || ||||| || || ||  | || ||||||
Db       2842 GTTGAAAGAGCCACCCTTGATTACAGCAAACACAGTTCTTTCCATTCTAGCAGTTGATTA 2901

Qy       1286 TCCTGTTGATAAGGTTTCTTGCTATGTTTCTGATGATGGTGCTGCAATGCTAACGTTTGA 1345
              |||||| ||||||||| | || ||||| ||  |  ||||||| || ||||| || |||||
```

```
Db       2902 TCCTGTGGATAAGGTTGCGTGTTATGTATCAAACAATGGTGCAGCTATGCTTACATTTGA 2961

Qy       1346 AGCATTATCTGAAACATCTGAATTTGCAAAGAAATGGGTTCCTTTCTGCAAACGGTACAA 1405
              |||  | ||||| ||| |||| ||||| |  |||||||||||||| || ||  ||  ||
Db       2962 AGCTCTCTCTGATACAGCTGAGTTTGCTAGAAAATGGGTTCCTTTTTGTAAGAAGTTTAA 3021

Qy       1406 TATTGAACCTCGCGCTCCAGAGTGGTACTTCCAACAGAAGATAGACTACTTGAAAGACAA 1465
              ||| || || || ||||| |||||||| ||    |||||||| || ||| |||| ||||
Db       3022 TATCGAGCCACGAGCTCCTGAGTGGTATTTTTCTCAGAAGATGGATTACCTGAAGAACAA 3081

Qy       1466 GGTGGCAGCAAACTTTGTTAGGGAGAGGAGAGCAATGA---------------------- 1503
                ||      |    ||||| |||||   |  | || ||||
Db       3082 AGTTCATCCTGCTTTTGTCAGGGAACGTCGTGCTATGAAGGTTTTCTTTGCTGCTTTTTC 3141

Qy       1504 ---------------------------------------------------------- 1503

Db       3142 TCTTTCTGAGTATATCCTATCATAAAAGTGTTGTTTCAAGAATCTGATTTACGTTTTTTG 3201

Qy       1504 ----------------AGAGAGAGTATGAGGAATTCAAGGTGAGAATCAATGCCTTAGTT 1547
                              ||||||| |||||||| || || ||||  || |||||  | |||
Db       3202 CTTGTTTGTTTGTTGCAGAGAGATTATGAGGAGTTTAAAGTGAAGATAAATGCACTGGTT 3261

Qy       1548 GCCAAAGCCCAGAAAGTTCCTGAAGAAGGATGGACAATGCAAGATGGAACCCCCTGGCCT 1607
              || |  || |||||||| ||||| ||||| ||||||||||||| || |||||||
Db       3262 GCTACTGCACAGAAAGTGCCTGAGGAAGGTTGGACTATGCAAGATGGAACTCCTTGGCCT 3321

Qy       1608 GGAAACAATGTTCGTGATCATCCTGGAATGATT---------------------------- 1640
              |||||||| || ||||| ||||||||||||||
Db       3322 GGAAACAACGTCCGTGACCATCCTGGAATGATTCAGGTAATGATGAGTTTGATTGAATAG 3381

Qy       1641 ---------------------------------------------------------- 1640

Db       3382 GCAAAAAAAAAGCGGTTTTTGTCCTCTTCACTTTGTTTCCCTGGATCTGTTAAATTGGAA 3441

Qy       1641 ---------------------------------------------------------- 1640

Db       3442 TGAGCACTCTACTTCTCAATATATCTTCAGACCGAAGCCTTTTTAAGAGATTTTGTAAAT 3501

Qy       1641 --CAGGTCTTCCTTGGCCAAAGCGGAGGCCTTGACTGTGAGGGAAATGAACTGCCACGAT 1698
                |||||  |||  | ||  ||  ||  ||||  | |||    ||  || |||||  | |||||
Db       3502 GACAGGTGTTCTTGGGTCATAGTGGAGTTCGTGATACGGATGGTAATGAGTTACCACGTC 3561

Qy       1699 TGGTTTATGTTTCTAGAGAGAAACGACCAGGCTATAACCATCATAAGAAAGCTGGTGCTA 1758
              | ||  |||||||| | ||||| || || || | | | || || |||||||||| ||||
Db       3562 TAGTGTATGTTTCTCGTGAGAAGCGGCCTGGATTTGATCACCACAAGAAAGCTGGAGCTA 3621

Qy       1759 TGAATGCATTGG------------------------------------------------ 1770
              |||||  | ||||
Db       3622 TGAATTCCTTGGTAAGTATAATGTGTTTCTTTATTTATGAATCTCTCTTTTCGGAGCCCT 3681

Qy       1771 ----------------------------------------TCCGAGTCTCTGCTGTAC 1788
                                                       |||||||||||||||| |
Db       3682 GACTTCTCATAAACTAAAACTCATCTTACTTCTTCTTGAAGATCCGAGTCTCTGCTGTTC 3741

Qy       1789 TAACAAATGCTCCATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGG 1848
              || |||| ||||| ||  |  | ||  | |||||||||||||||||||||||||||| |
Db       3742 TATCAAACGCTCCTTACCTTCTTAATGTCGATTGTGATCACTACATCAACAACAGCAAAG 3801

Qy       1849 CTATAAAGGAAGCAATGTGTTTTATGATGGACCCTTTACTAGGAAAGAAGGTTTGCTATG 1908
              | || |  ||| | |||||||| |||||||||||  |  |||||||| ||||| ||||
Db       3802 CAATTAGAGAATCTATGTGTTTCATGATGGACCCGCAATCGGGAAAGAAAGTTTGTTATG 3861

Qy       1909 TACAGTTCCCTCAAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATG 1968
              | ||||| || || |||||||||||||||||| | |||||  ||||  | |||||  || |
Db       3862 TTCAGTTTCCGCAGAGATTTGATGGGATTGATAGACATGATAGATACTCAAACCGTAACG 3921

Qy       1969 TTGTCTTTTTTGAT---------------------------------------------- 1982
              ||||  || ||||||
Db       3922 TTGTGTTCTTTGATGTATGTGTCCTTATCTCTTTTGCTTTGTTTCTGTTTATGTTTTAGT 3981

Qy       1983 ----------------------------------------------ATCAACATGAAAGGT 1997
                                                           ||  ||||||||||||
Db       3982 GCTTTTCCTCTTTTCTCATTTGATATTGTTTTGGTGTGGAAACAGATTAACATGAAAGGT 4041
```

```
Qy       1998 TTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTGTATTTAGAAGGCAGGCA 2057
Db       4042 CTTGATGGGATACAAGGACCGATATATGTCGGGACAGGTTGTGTGTTTAGAAAACAGGCT 4101

Qy       2058 TTATATGGTTATGATGCCCCCAAAACAAAGAAGCCACCATCAAGGACTTGCAACTGCTGG 2117
Db       4102 CTTTATGGTTTTGATGCACCAAAGAAGAAGAAACCACCAGGCAAAACCTGTAACTGTTGG 4161

Qy       2118 CCCAAGTGGTGCTTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAA 2177
Db       4162 CCTAAATGGTGTTGTTTGTGTTGTGGGTTGAGAAAGAAGAGTAAAACGAAAGCCAAAGAT 4221

Qy       2178 CCCAAAACAGAGAAGAAAAAGTTATTATTTTTC---------------------------- 2210
Db       4222 AAGAAAACTAACACTAAAGAGACTTCAAAGCAGATTCATGCGCTAGAGAATGTCGACGAA 4281

Qy       2211 --------------------------AAGAAAGAAGAGAACCAATCCCCTGCATATGCTC 2244
Db       4282 GGTGTTATCGTCCCAGGTAAAAAAAGAAGGAAAAAAAAAAACATTTCTTATTTGGTTTCTG 4341

Qy       2245 TTGGTGAAATTGACGAAGCTGCTCC-----------AGGAGCTGAGAATGAAAAGGCCGG 2293
Db       4342 TCTTGTTGAAAGTCTAAGTAGATCCTTTTGATTGTTAGTGTCAAATGTTGAGAAGAGATC 4401

Qy       2294 TATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTACATC 2353
Db       4402 TGAAGCAACACAATTGAAATTGGAGAAGAAGTTTGGACAATCTCCGGTTTTCGTTGCCTC 4461

Qy       2354 CACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAAAGA 2413
Db       4462 TGCTGTTCTACAGAACGGTGGAGTTCCCCGTAACGCAAGCCCCGCATGTTTGTTAAGAGA 4521

Qy       2414 AGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAG------- 2466
Db       4522 AGCCATTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGAATGGGGAAAAGAGGTAGA 4581

Qy       2467 ------------------------------------------------------------ 2466
Db       4582 AAACATTACAAAGTTTTTCAACTTCTGAAAACTAGAAAAGTTCTTGTGATCTCATTCTTG 4641

Qy       2467 ---------------AGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTCTAACTG 2511
Db       4642 CTGATAATCACACGCAGATCGGGTGGATTTATGGATCGGTGACTGAAGATATCCTGACGG 4701

Qy       2512 GTTTCAAGATGCATTGTCATGGTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCAT 2571
Db       4702 GTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGCCTAAGCGTGCAGCTT 4761

Qy       2572 TCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTC 2631
Db       4762 TTAAAGGATCTGCTCCTATTAACTTGTCAGATCGTCTTCATCAAGTTCTACGTTGGGCTC 4821

Qy       2632 TTGGGTCTATTGAGATCTTCTTCAGCAATCATTGCCCTCTTTGGTATGGGTATGGTGGCG 2691
Db       4822 TTGGCTCTGTAGAGATTTTCTTGAGCAGACATTGTCCGATATGGTATGGTTATGGTGGTG 4881

Qy       2692 GTCTGAAATTTTTGGAAAGATTTTCCTACATCAACTCCATCGTGTATCCTTGGACATCTA 2751
Db       4882 GTTTAAAATGGTTGGAGAGATTCTCTTACATCAACTCTGTCGTCTATCCTTGGACTTCAC 4941

Qy       2752 TTCCCCTCTTGGCTTACTGTACATTGCCTGCCATCTGTTTATTGACAGGGAAATTTATCA 2811
Db       4942 TTCCATTGATCGTCTATTGTTCTCTCCCCGCGGTTTGTTTACTCACAGGAAAATTCATCG 5001

Qy       2812 CTCCAGAG 2819
Db       5002 TCCCTGAG 5009
```

RESULT 11
US-09-313-294A-849
; Sequence 849, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 849
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549940H1
US-09-313-294A-849

Query Match 6.8%; Score 268; DB 4; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.1e-66;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2849 GTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGG 2908
Db          1 GTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAGTGGTGTTGG 60

Qy       2909 AATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCT 2968
Db         61 AATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTCCTCACACCT 120

Qy       2969 CTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGT 3028
Db        121 CTTTGCTGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTTGATACAAGCTTCACCGT 180

Qy       3029 GACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTAC 3088
Db        181 GACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGGACTAC 240

Qy       3089 CTTATTGATACCTCCTACCACCTTGCTT 3116
Db        241 CTTATTGATACCTCCTACCACCTTGCTT 268
```

RESULT 12
US-09-313-294A-7123
; Sequence 7123, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7123
; LENGTH: 317
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381163H1
; NAME/KEY: unsure
; LOCATION: 10
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7123

Query Match 5.4%; Score 213; DB 4; Length 317;
Best Local Similarity 83.3%; Pred. No. 1.5e-50;

```
  Matches  265;  Conservative    0;  Mismatches   51;  Indels    2;  Gaps    2;

Qy       2172 ACCAAACCCAAAACAGAGAAGAAAAAGTTATTATTTTTCAAGAAAGAAGAGAACCAATCC 2231
             || |||||  | || |||||||| ||   ||||||||||||||||| ||| |||| |||
Db          1 ACAAAACCANAGACGGAGAAGAAGAAAAGATTATTTTTCAAGAAAGCAGAAAACCCATCT 60

Qy       2232 -CCTGCATATGCTCTTGGTGAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGC 2290
              |||||||||||| | |||||||||||| |||| ||||||||| ||||| |  || |||||
Db         61 ACCTGCATATGCTTTGGGTGAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAGGC 120

Qy       2291 CGGTATTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTTTGTTAC 2350
              ||| || |||||||||||| ||| |||| |||||||||| || ||||||||||||||  |
Db        121 CGGAATCGTAAATCAACAGAAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTTTGTCGC 180

Qy       2351 ATCCACACTTCTCGAGAATGGTGGAACCTTGAAGAGTGCAAGTCCTGCTTCTCTTTTGAA 2410
              ||| |||||||| ||||| || || ||| ||||||| |||||||| |||||||| ||||
Db        181 ATCAACACTTCTTGAGAACGGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTCTGAA 240

Qy       2411 AGAAGCTATACATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGAT 2470
              |||||||||||||| || || || || || |||||||||| ||||||||| ||||||
Db        241 GGAAGCTATACATGTTATCAGCTGCGGCTACGAAGACAAGACCGACTGGGGATAAGAGA- 299

Qy       2471 TGGCTGGATCTATGGATC 2488
              ||||||||| || |||||
Db        300 TGGCTGGATTTACGGATC 317


RESULT 13
US-09-313-294A-4015
; Sequence 4015, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 4015
;   LENGTH: 278
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700347541H1
;   NAME/KEY: unsure
;   LOCATION: 32, 38, 125, 162, 236, 246, 248
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4015

  Query Match             4.2%;  Score 164.2;  DB 4;  Length 278;
  Best Local Similarity   87.4%;  Pred. No. 1.1e-36;
  Matches  201;  Conservative    0;  Mismatches   26;  Indels    3;  Gaps    2;

Qy       2363 CGAGAATGGTGGAACCTTGAAGAGTGCAAGT-CCTGCTTCTCTTTTGAAAGAAGCTATAC 2421
              ||| |||||||||||||||||||||||||| ||||| |||||| ||||||||||||||||
Db          1 CGAAAATGGTGGAACCTTGAAGAGTGCAAGTNCCTGCNTCTCTTCTGAAAGAAGCTATAC 60

Qy       2422 ATGTCATTAGTTGTGGTTATGAAGACAAGACAGACTGGGGAAAAGAGATTGGCTGGATCT 2481
              |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db         61 ATGTCATTAGTTGTGGTTATGAAGACAAGACTGATTGGGGAAAAGAGATTGGCTGGATCT 120

Qy       2482 ATGGATCAGTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTCATGGTTGGC--G 2539
              |||| |||||||||| ||||||||||||||||||||||| ||      |||||||  |
Db        121 ATGGNTCAGTTACAGAGGATATTCTAACTGGTTTCAAGATGNATGTCATGGGTTGGCCGG 180

Qy       2540 GTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTC 2589
                | |||||||||||||||||| |   ||||||||||    ||||
Db        181 TCAATTTTACTGCATACCTAAACGGCTGCAATTCAAAGGTTCGCAGCCTC 230


RESULT 14
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 6485
;   LENGTH: 281
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
US-09-313-294A-6485

  Query Match             4.0%;  Score 159;  DB 4;  Length 281;
  Best Local Similarity   73.1%;  Pred. No. 3.2e-35;
  Matches  204;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy       2720 CATCAACTCCATCGTGTATCCTTGGACATCTATTCCCCTCTTGGCTTACTGTACATTGCC 2779
              ||||||| ||| | | || ||    || ||  | || ||| |   |||||||||   ||||
Db          1 CATCAACACCACCATCTACCCGCTCACGTCCCTCCCGCTCCTCATTTACTGTATCCTGCC 60

Qy       2780 TGCCATCTGTTTATTGACAGGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCT 2839
              |||||||||  |  | || ||||| || ||||  ||||||  | |  ||  | || ||  |
Db         61 TGCCATCTGCCTGCTCACGGGGAAGTTCATCATCCCAGAGATCAGCAACTTCGCTAGTAT 120

Qy       2840 GTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGAGCATCCTAGAAATGAGATGGAG 2899
              |||||||| || || || ||||  || || || ||| | ||||| || ||||| |||||
Db        121 CTGGTTCATCTCTCTCTTCATCTCGATCTTCGCCACGGGTATCCTGGAGATGAGGTGGAG 180

Qy       2900 TGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTGTC 2959
              || || || || || || |||||||||| |||||||||||||||| |||||  | ||
Db        181 CGGCGTGGGCATCGACGAGTGGTGGAGGAACGAGCAGTTCTGGGTCATCGGAGGCATCTC 240

Qy       2960 CTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGT 2998
              | | |||||||| || || |||||||| || ||||||||
Db        241 CGCCCACCTCTTCGCCGTCTTCCAGGGCCTCCTCAAGGT 279


RESULT 15
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 4753
;   LENGTH: 300
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
;   NAME/KEY: unsure
;   LOCATION: 64, 189
;   OTHER INFORMATION: a, t, c, g, or other
```

US-09-313-294A-4753

Query Match 3.7%; Score 145; DB 4; Length 300;
Best Local Similarity 70.0%; Pred. No. 3.2e-31;
Matches 208; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

```
Qy       2818 AGCTGAATAATGTTGCCAGCCTGTGGTTCATGTCACTTTTTATCTGCATTTTTGCTACGA 2877
             || | |  ||  | |||||| | |||||||| || || || ||||  || || || |||
Db          3 AGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATCTCGATCTTCGCCACGG 62

Qy       2878 GCATCCTAGAAATGAGATGGAGTGGTGTTGGAATTGATGACTGGTGGAGGAATGAGCAGT 2937
             | ||||| || ||||| ||||| || || || || || || |||||||||||| |||||||
Db         63 GNATCCTGGAGATGAGGTGGAGCGGGGTGGGCATCGACGAGTGGTGGAGGAACGAGCAGT 122

Qy       2938 TCTGGGTCATTGGAGGTGTGTCCTCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGG 2997
             ||||||| || || ||  | ||| | |||||||| || |||||||||||| || |||||||
Db        123 TCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTCGCCGTGTTCCAGGGCCTGCTCAAGG 182

Qy       2998 TCATAGCTGGTGTTGATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGT 3057
             |  | |  ||  | || || | |||||||||| || || ||||     || || || |
Db        183 TGCTGGNCGGCATCGACACCAACTTCACCGTCACCTCCAAGGCCTCGGACGAGGACGGCG 242

Qy       3058 TCTCAGAG--CTATATACATTCAAATGGACTACCTTATTGATACCTCCTACCACCTT 3112
               ||  | |  || || |  ||||| ||||| ||  |  | || || || |||||| |
Db        243 ACTTCGCGGACTGTACATGTTCAAGTGGACGACGCTCCTTATCCCGCCCACCACCAT 299
```

Search completed: December 16, 2003, 00:53:10
Job time : 248 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:00:59 ; Search time 27 Seconds
(without alignments)
3868.118 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4278.5	72.3	1084	2	T08583	cellulose synthase
2	4209.5	71.1	1088	2	H84604	probable cellulose
3	4202	71.0	1081	2	T52028	cellulose synthase
4	3967.5	67.0	1026	2	T51579	cellulose synthase
5	3945.5	66.7	1081	2	T05351	cellulose synthase
6	3875.5	65.5	1065	2	T52054	cellulose synthase
7	3782	63.9	1065	2	F84649	probable cellulose
8	3469.5	58.6	974	2	T10797	cellulose synthase
9	3329.5	56.3	958	2	T04870	cellulose synthase
10	2907	49.1	685	2	T10800	cellulose synthase
11	2330.5	39.4	1181	2	D86157	hypothetical prote
12	2311	39.1	1145	2	T51546	cellulose synthase
13	2256	38.1	1111	2	T05646	hypothetical prote
14	2166.5	36.6	1036	2	D84741	probable cellulose
15	2118.5	35.8	583	2	T02209	cellulose synthase
16	2114	35.7	979	2	C86446	probable cellulose
17	1056.5	17.9	757	2	T02561	probable cellulose
18	1046	17.7	712	2	T02552	cellulose synthase
19	1040	17.6	755	2	T02553	cellulose synthase
20	993.5	16.8	748	2	T02560	cellulose synthase
21	959.5	16.2	689	2	T08918	hypothetical prote
22	952	16.1	727	2	T08920	hypothetical prote
23	944.5	16.0	686	2	T08919	hypothetical prote
24	901	15.2	828	2	E71417	hypothetical prote
25	829	14.0	710	2	B71417	hypothetical prote
26	414.5	7.0	326	2	T08591	TGACG-motif bindin
27	402.5	6.8	693	2	AF2275	cellulose synthase
28	393	6.6	322	2	T12093	TGACG-motif bindin
29	359.5	6.1	322	2	T08592	TGACG-motif-bindin
30	256.5	4.3	759	2	D70422	cellulose synthase
31	252	4.3	754	2	A43735	bcsA protein - Ace
32	248	4.2	874	2	AB0985	probable polysacch
33	239.5	4.0	1596	2	T31338	cellulose synthase
34	239	4.0	888	2	E91180	probable cellulose
35	238	4.0	692	2	S47754	hypothetical prote
36	238	4.0	888	2	H65151	hypothetical 101.6
37	232	3.9	888	2	G86026	probable cellulose
38	227.5	3.8	729	2	A98320	cellulose synthase
39	227.5	3.8	729	2	AD2963	cellulose synthase
40	227.5	3.8	861	2	I39714	cellulose synthase
41	211	3.6	664	2	A95889	probable cellulose
42	201	3.4	723	2	C36963	cellulose synthase
43	194	3.3	768	2	B97083	glycosyltransferas
44	175	3.0	654	2	AB2932	beta 1,3 glucan sy
45	175	3.0	654	2	D98350	hypothetical prote

ALIGNMENTS

RESULT 1
T08583
cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana
N;Alternate names: protein T22F8.250
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C;Accession: T08583; T09014
R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; M
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16442
A;Accession: T08583
A;Molecule type: DNA
A;Residues: 1-1084 <BEV>
A;Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.250
A;Experimental source: cultivar Columbia; BAC clone T22F8
R;Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.;
Science 279, 717-720, 1998
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Reference number: Z13745; MUID:98111412; PMID:9445479
A;Accession: T09014
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1084 <ARI>
A;Cross-references: EMBL:AF027173; NID:g2827140; PIDN:AAC39335.1; PID:g2827141
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:T22F8.250; Ath-A
A;Map position: 4
A;Introns: 27/3; 156/3; 191/1; 254/2; 343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3;
C;Keywords: cell wall synthesis; glycosyltransferase; hexosyltransferase

Query Match 72.3%; Score 4278.5; DB 2; Length 1084;
Best Local Similarity 71.3%; Pred. No. 0;
Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:|||||||| |:|  |      :  :| :|| ||||||:: |    : ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
              ||||||| ||||||||| | ||||||||||:||  || | ||||: :|||: ||:  | |
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEFD-HGMD 119

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
               :  ||: |   :: |||| :  |        :|||      |:  ::|||:     |
Db        120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
               | |::| |: | | | | ||| | ||:| |||||||||:||| ||:|| |::    |:
Db        175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy        235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290
```

```
Db        235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy        291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350

Db        295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy        351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410

Db        355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSDTAEFA 414

Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEPKVRINALVAK 470

Db        415 RKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530

Db        475 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590

Db        535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650

Db        595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKKP 654

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEA--A 708

Db        655 PGKTCNCWPKWC-CLCC-GLRKKSKT---KAKDKKT---NTKETSKQIHALENVDEGVIV 706

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768

Db        707 PVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGYED 766

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828

Db        767 KTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRW 826

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888

Db        827 ALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTGKF 886

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948

Db        887 IVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALFQG 946

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008

Db        947 LLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAISN 1006

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068

Db       1007 GYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRVNP 1066

Qy       1069 FLAKDDGPLLEECGLDC 1085

Db       1067 FVAK-GGPVLEICGLNC 1082


RESULT 2
H84604
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84604
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1088 <STO>
A;Cross-references: GB:AE002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g21770
A;Map position: 2

  Query Match             71.1%;  Score 4209.5;  DB 2;  Length 1088;
  Best Local Similarity   70.2%;  Pred. No. 1.7e-314;
  Matches  769;  Conservative  127;  Mismatches  179;  Indels   21;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db          1 MNTGGRLIAGSHNRNEFVLINADDTARIRSAEELSGQTCKICRDEIELTDNGEPFIACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db         61 CAFPTCRPCYEYERREGNQACPQCGTRYKRIKGSPRVEGDEEDDDIDDLEHEFY--GMDP 118

Qy        121 QSVAESML-YGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179

Db        119 EHVTEAALYYMRLNTGRGTDEVSHLYSASPGSEVPLLTYCDEDSDMYSDRHALIVPPSTG 178

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234

Db        179 LGNRVHHVPFTDSFASIHTRPMVPQKDLTVYGYGSVAWKDRMEVWKKQQIEKLQVVKNER 238

Qy        235 GNDGGG-----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHY 289

Db        239 VNDGDGDGFIVDELDDPGLPMMDEGRQPLSRKLPIRSSRINPYRMLIFCRLAILGLFFHY 298

Qy        290 RVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQL 349

Db        299 RILHPVNDAFGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSEL 358

Qy        350 APIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEF 409

Db        359 APVDVFVSTVDPLKEPPLITANTVLSILAVDYPVEKVACYVSDDGAAMLTFEALSYTAEF 418

Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469

Db        419 ARKWVPFCKKFSIEPRAPEWYFSQKMDYLKHKVDPAFVMERRAMKRDYEEFKVKINALVS 478

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529

Db        479 VSQKVPEDGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVCDMDGNELPRLVYVSREKRPG 538

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589

Db        539 FDHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIREAMCFMMDPQSGKKICYV 598

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649

Db        599 QFPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKQ 658

Qy        650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAP 709

Db        659 PPGRTCNCWPKWC-CLCC-GMRK-KKTGKVKDNQRK----KPKETSKQIHALEHIEEGLQ 711

Qy        710 GAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDK 769

Db        712 VTNAENNSETAQLKLEKKFGQSPVLVASTLLLNGGVPSNVNPASLLRESIQVISCGYEEK 771

Qy        770 TDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWA 829

Db        772 TEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRWA 831

Qy        830 LGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI 889

Db        832 LGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLLVYCSLPAICLLTGKFI 891

Qy        890 TPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGL 949
```

```
Db        892 VPEISNYAGILFLLMFMSIAVTGILEMQWGKIGIDDWWRNEQFWVIGGVSSHLFALFQGL 951

Qy        950 LKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNG 1009

Db        952 LKVLAGVSTNFTVTSKAADDGEFSELYIFKWTSLLIPPTTLLIINIVGVIVGVSDAINNG 1011

Qy       1010 YESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPF 1069

Db       1012 YDSWGPLFGRLFFALWVIVHLYPFLKGLLGKQDRVPTIILVWSILLASILTLLWVRVNPF 1071

Qy       1070 LAKDDGPLLEECGLDC 1085

Db       1072 VSK-DGPVLEICGLDC 1086


RESULT 3
T52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52028
R;Joshi, C.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z25890
A;Accession: T52028
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1081 <JOS>
A;Cross-references: EMBL:AF062485; PIDN:AAC29067.1

  Query Match           71.0%;  Score 4202;  DB 2;  Length 1081;
  Best Local Similarity 70.8%;  Pred. No. 6.5e-314;
  Matches  770;  Conservative  126;  Mismatches  172;  Indels   20;  Gaps   10;

Qy          7 LVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVC 66

Db          4 LIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNECAFPVC 63

Qy         67 RDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW--DGHDSQSVA 124

Db         64 RPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDDIDDLDNEFEYGNNGIGFDQVS 123

Qy        125 ESMLYGHMSYG-RGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL-VPSFMGGGGK 182

Db        124 EGMSISRRNSGFPQSDLDSAPPGSQ----IPLLTYGDEDVEISSDRHALIVPPSLGGHGN 179

Qy        183 RIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERMHQTGNDGGGD 241

Db        180 RVHPVSLSDPTVAAHRRLMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVRHEGDPD 239

Qy        242 --DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAF 299

Db        240 FEDGDDADFPMMDEGRQPLSMKIPIKSSKINPYRMLIVLRLVILGLFFHYRILHPVKDAY 299

Qy        300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTV 359

Db        300 ALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVDVFVSTV 359

Qy        360 DPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 419

Db        360 DPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSETAEFARKWVPFCKK 419

Qy        420 YNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 479

Db        420 YCIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQKVPEDGW 479

Qy        480 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAM 539

Db        480 TMQDGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPGFDHHKKAGAM 539

Qy        540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 599
```

```
Db        540 NSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRFDGID 599

Qy        600 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 659

Db        600 RHDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKGPRKTCNCWP 659

Qy        660 KWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA--PGAENEKAG 717

Db        660 KWCL--LCFGSRKNRKAKTVAADKKK----KNREASKQIHALENIEEGRGHKVLNVEQST 713

Qy        718 IVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIG 777

Db        714 EAMQMKLQKKYGQSPVFVASARLENGGMARNASPACLLKEAIQVISRGYEDKTEWGKEIG 773

Qy        778 WIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFF 837

Db        774 WIYGSVTEDILTGSKMHSHGWRHVYCTPKLAAFKGSAPINLSDRLHQVLRWALGSVEIFL 833

Qy        838 SNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVA 897

Db        834 SRHCPIWYGYGGGLKWLERLSYINSVVYPWTSLPLIVYCSLPAICLLTGKFIVPEISNYA 893

Qy        898 SLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVD 957

Db        894 SILFMALFSSIAITGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLAGVD 953

Qy        958 TSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLF 1017

Db        954 TNFTVTSKAADDGEFSDLYLFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYDSWGPLF 1013

Qy       1018 GKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPL 1077

Db       1014 GRLFFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVAK-GGPI 1072

Qy       1078 LEECGLDC 1085

Db       1073 LEICGLDC 1080


RESULT 4
T51579
cellulose synthase catalytic subunit (IRX3) - Arabidopsis thaliana
N;Alternate names: protein T10B6_80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C;Accession: T51579
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; M
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51579
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1026 <SAT>
A;Cross-references: EMBL:AL391142
A;Experimental source: cultivar Columbia; BAC clone T10B6
C;Genetics:
A;Map position: 5
A;Introns: 25/3; 91/1; 184/1; 215/2; 304/2; 419/3; 507/3; 578/3; 650/1; 715/3; 833/3
A;Note: T10B6_80

  Query Match           67.0%;  Score 3967.5;  DB 2;  Length 1026;
  Best Local Similarity 66.8%;  Pred. No. 6.1e-296;
  Matches  736;  Conservative  129;  Mismatches  145;  Indels   91;  Gaps   16;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db          1 MEASAGLVAGSHNRNELVVIHNHEE--PKPLKNLDGQFCEICGDQIGLTVEGDLFVACNE 58

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db         59 CGFPACRPCYEYERREGTQNCPQCKTRYKRLRGSPRVEGDEDEEDIDDIEYEFNIE-HEQ 117

Qy        121 ---QSVAESMLYGHMSYGRG--GDPNGA-PQAFQLNPNVPLLTNGQMVDDIPPEQHALVP 174
```

```
Db        118 DKHKHSAEAMLYGKMSYGRGPEDDENGRFP---------PVIAGGHSGE---------- 157

Qy        175 SFMGGGG---------KRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWK 225

Db        158 -FPVGGGYGNGEHGLHKRVHPY--------------PSSEAGSEG----GWRERMDDWK 197

Qy        226 QRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285

Db        198 L------QHGNLGPEPD-DDPEMGLIDEARQPLSRKVPIASSKINPYRMVIVARLVILAV 250

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQ 345

Db        251 FLRYRLLNPVHDALGLWLTSVICEIWFAVSWILDQFPKWFPIERETYLDRLSLRYEREGE 310

Qy        346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405

Db        311 PNMLAPVDVFVSTVDPLKEPPLVTSNTVLSILAMDYPVEKISCYVSDDGASMLTFESLSE 370

Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465

Db        371 TAEFARKWVPFCKKFSIEPRAPEMYFTLKVDYLQDKVHPTFVKERRAMKREYEEFKVRIN 430

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525

Db        431 AQVAKASKVPLEGWIMQDGTPWPGNNTKDHPGMIQVFLGHSGGFDVEGHELPRLVYVSRE 490

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585

Db        491 KRPGFQHHKKAGAMNALVRVAGVLTNAPFMLNLDCDHYVNNSKAVREAMCFLMDPQIGKK 550

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645

Db        551 VCYVQFPQRFDGIDTNDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFKRQALYGYEPP 610

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705

Db        611 KGPKRPKMIS------CGCCPCFGRRRKNKK-----------FSKNDMNGDVAALG--- 649

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765

Db        650 ----GAEGDKEHLMSEMNFEKTFGQSSIFVTSTLMEEGGVPPSSSPAVLLKEAIHVISCG 705

Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825

Db        706 YEDKTEWGTELGWIYGSITEDILTGFKMHCRGWRSIYCMPKRPAFKGSAPINLSDRLNQV 765

Qy        826 LRWALGSIEIFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLL 884

Db        766 LRWALGSVEIFFSRHSPLWYGYKGGKLKWLERFAYANTTIYPFTSIPLLAYCILPAICLL 825

Qy        885 TGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFA 944

Db        826 TDKFIMPPISTFASLFFISLFMSIIVTGILELRWSGVSIEEWWRNEQFWVIGGISAHLFA 885

Qy        945 VFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004

Db        886 VVQGLLKILAGIDTNFTVTSKATDDDDFGELYAFKWTTLLIPPTTVLIINIVGVVAGISD 945

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064

Db        946 AINNGYQSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPTIVVIWSVLLASIFSLLWV 1005

Qy       1065 RIDPFLAKDDGPLLEECGLDC 1085

Db       1006 RIDPFVLKTKGPDTSKCGINC 1026


RESULT 5
T05351
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana
N;Alternate names: protein F8B4.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T05351
R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; D
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15409
A;Accession: T05351
A;Molecule type: DNA
A;Residues: 1-1081 <BEV>
A;Cross-references: EMBL:AL034567
A;Experimental source: cultivar Columbia; BAC clone F8B4
C;Genetics:
A;Gene: RSW1
A;Map position: 4
A;Introns: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1;
A;Note: F8B4.110
C;Keywords: glycosyltransferase; hexosyltransferase; P-loop

  Query Match             66.7%;  Score 3945.5;  DB 2;  Length 1081;
  Best Local Similarity   66.9%;  Pred. No. 3.2e-294;
  Matches  727;  Conservative  140;  Mismatches  182;  Indels   37;  Gaps   11;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db          1 MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMNGQICQICGDDVGLAETGDVFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db         61 CAFPVCRPCYEYERKDGTQCCPQCKTRFRRHRGSPRVEGDEDEDDVDDIENEFNY----A 116

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDD---IPPEQHALVPSF- 176

Db        117 QGANKA------RHQRHGEEFSSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER--MHQT 234

Db        170 LGPSDRNAISSPYIDPRQPVPVRIVDPSKDLNSYGLGNVDWKERVEGWKLKQEKNMLQMT 229

Qy        235 GN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285

Db        230 GKYHEGKGGEIEGTGSNGE--ELQMADDTRLPMSRVVPIPSSRLTPYRVVIILRLIILCF 287

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDKEGQ 345

Db        288 FLQYRTTHPVKNAYPLWLTSVICEIWFAFSWLLDQFPKWYPINRETYLDRLAIRYDRDGE 347

Qy        346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405

Db        348 PSQLVPVDVFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVACYVSDDGSAMLTFESLSE 407

Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465

Db        408 TAEFAKKWVPFCKKFNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKREYEEFKVRIN 467

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525

Db        468 ALVAKAQKIPEEGWTMQDGTPWPGNNTRDHPGMIQVFLGHSGGLDTDGNELPRLIYVSRE 527

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585

Db        528 KRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK 587

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645

Db        588 CCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFNRQALYGYDPV 647

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705

Db        648 LTEE------DLEPNIIVKSCCGSRKKGKSSKKYNYEKRRGI--NRSDSNAPLFNMEDID 699

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765

Db        700 EGFEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIPPTTNPATLLKEAIHVISCG 759
```

```
Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825
              |||||:||||||||||||||||||||||  || |||| | | |||||||:||||||:||
Db        760 YEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCNPPRPAFKGSAPINLSDRLNQV 819

Qy        826 LRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLT 885
              ||||||||||  | |||:|||| | |: ||| :|||:|||| |||||:||| ||| ||:|
Db        820 LRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINTIVYPITSIPLIAYCILPAFCLIT 879

Qy        886 GKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAV 945
               :|| ||::| ||:||: ||| |  | |||:||||| |:|||||||||||| |:|||||
Db        880 DRFIIPEISNYASIWFILLFISIAVTGILELRWSGVSIEDWWRNEQFWVIGGTSAHLFAV 939

Qy        946 FQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004
              |||||||:||:||:||||||  | | :|:||| |||| |||||||:||:| ||:|||||
Db        940 FQGLLKVLAGIDTNFTVTSKATDEDGDFAELYIFKWTALLIPPTTVLLVNLIGIVAGVSY 999

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
              |:|:||:|||||||||||| ||| |||||||||:|||||||||||||:||||||||||
Db       1000 AVNSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPTIVIVWSVLLASIFSLLWV 1059

Qy       1065 RIDPFL 1070
              ||:||:
Db       1060 RINPFV 1065


RESULT 6
T52054
cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H
Science 279, 717-720, 1998
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Reference number: Z13745; MUID:98111412; PMID:9445479
A;Accession: T52054
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN:AAC39336.1
C;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; involved in as
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             65.5%;  Score 3875.5;  DB 2;  Length 1065;
  Best Local Similarity   66.7%;  Pred. No. 7.4e-289;
  Matches  728;  Conservative  145;  Mismatches  161;  Indels   57;  Gaps   18;

Qy         23 DGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCP 82
              :|:   || :    | |||| |:||    || ||||: |:||||| ||||||::| |:||
Db          4 EGETAGKPMKNIVPQTCQICSDNVGKTVDGDRFVACDICSFPVCRPCYEYERKDGNQSCP 63

Qy         83 QCKTRYKRLKGCQRVTGDEEEDGV-DDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPN 141
              |||||||||||   : ||::|||: |:   |||:     : ::| ||  |:: |: |:
Db         64 QCKTRYKRLKGSPAIPGDKDEDGLADEGTVEFNYP--QKEKISERMLGWHLTRGK-GEEM 120

Qy        142 GAPQAFQ--LNPNVPLLTNGQMVDDIPPEQHALVPSFMG-----GGGKRIHPLPY-ADPS 193
              | ||  :   : ::| ||: |   |   |  |  |  :       ||||   ||| :| :
Db        121 GEPQYDKEVSHNHLPRLTSRQ---DTSGEFSAASPERLSVSSTIAGGKR---LPYSSDVN 174

Qy        194 LPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER-----MHQTGNDGGGDDGDD--- 245
                  | :||      | |:||||||:: || :||:       |  :: || | |
Db        175 QSPNRRIVDP------VGLGNVAWKERVDGWKMKQEKNTGPVSTQAASERGGVDIDASTD 228

Qy        246 --ADLPLM-DEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALW 302
                ||  |: ||||| |||: :|||:||||||:|::|||:|  | |||: :|| :|||||
Db        229 ILADEALLNDEARQPLSRKVSIPSSRINPYRMVIMLRLVILCLFLHYRITNPVPNAFALW 288

Qy        303 LISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPL 362
              |:|||||||||:|||||||||||: ||||||||:||:|:||:||||| :| ||||||||
```

```
Db        289 LVSVICEIWFALSWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFVSTVDPL 348

Qy        363 KEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNI 422
              ||||||| |||||||:|||||||||||| ||||||||:||:|:|||||||:|||||||:|:|
Db        349 KEPPLVTANTVLSILAVDYPVDKVSCYVFDDGAAMLSFESLAETSEFARKWVPFCKKYSI 408

Qy        423 EPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQ 482
              |||||||| |||||||| :||::|||||||||||:||||||:|| | ||||| ||
Db        409 EPRAPEWYFAAKIDYLKDKVQTSFVKDRRAMKREYEEFKIRINALVSKALKCPEEGWVMQ 468

Qy        483 DGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNAL 542
              |||||||||  |||||||||:|||| |||||||||||||||: ||||||||||
Db        469 DGTPWPGNNTGDHPGMIQVFLGQNGGLDAEGNELPRLVYVSREKRPGFQHHKKAGAMNAL 528

Qy        543 VRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHD 602
              |||||||| |::|||||||||||||::|||||:||| |||:|||||||||||::|
Db        529 VRVSAVLTNGPFILNLDCDHYINNSKALREAMCFLMDPNLGKQVCYVQFPQRFDGIDKND 588

Qy        603 RYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP---KTKKPP--SRTCNC 657
              |||||| ||||||::|||||||:|||||||| | |||||: |   | |||   |: |
Db        589 RYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPIKVKHKKPSLLSKLCG- 647

Qy        658 WPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGA--ENEK 715
                          |:||:    | :::||| :  ::  | : | :|:|   ||  ::||
Db        648 ----------GSRKKNSKAKKESDKKK--SGRHTDSTVPVFNLDDIEEGVEGAGFDDEK 694

Qy        716 AGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKE 775
              | :::|  |||:||||:||| |||:||||   ||:| :|||||||||||||:||| |
Db        695 ALLMSQMSLEKRFGQSAVFVASTLMENGGVPPSATPENLLKEAIHVISCGYEDKSDWGME 754

Qy        776 IGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEI 835
              |||||||||||||||||  |||||||:||  |||||||:|||||||:||||||||:||
Db        755 IGWIYGSVTEDILTGFKMHARGWRSIYCMPKLPAFKGSAPINLSDRLNQVLRWALGSVEI 814

Qy        836 FFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNN 895
              || |||:|||| | |||||||:|:|: :|| |||||| |||| |:|| | :|| |:::|
Db        815 LFSRHCPIWYGYNGRLKFLERFAYVNTTIYPITSIPLLMYCTLLAVCLFTNQFIIPQISN 874

Qy        896 VASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAG 955
              :||:||:|||: |||| |||||||||||:|||||||||||||||:||||||||:|||:||
Db        875 IASIWFLSLFLSIFATGILEMRWSGVGIDEWWRNEQFWVIGGVSAHLFAVFQGILKVLAG 934

Qy        956 VDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWG 1014
              :||:||||||  | | :|:||| |||||||||||||::| :|||||| |||:||:|||
Db        935 IDTNFTVTSKASDEDGDFAELYLFKWTTLLIPPTTLLIVNLVGVVAGVSYAINSGYQSWG 994

Qy       1015 PLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDD 1074
              |||||||||||||||||||||:|||||||||:|||:|||||||||||||||| ::
Db        995 PLFGKLFFAFWVIVHLYPFLKGLMGRQNRTPTIVVVWSVLLASIFSLLWVRIDPFTSRVT 1054

Qy       1075 GPLLEECGLDC 1085
              || : |||::|
Db       1055 GPDILECGINC 1065


RESULT 7
F84649
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84649
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84649
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1065 <STO>
A;Cross-references: GB:AE002093; NID:g4432865; PIDN:AAD20713.1; GSPDB:GN00139
C;Genetics:
```

```
A;Gene: At2g25540
A;Map position: 2

  Query Match             63.9%;  Score 3782;  DB 2;  Length 1065;
  Best Local Similarity   63.9%;  Pred. No. 1.1e-281;
  Matches  702;  Conservative  150;  Mismatches  168;  Indels   78;  Gaps   15;

Qy          7 LVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVC 66
              :||||: | | |  | | | | || :: |||:||||||||||   |: |||||| ||:|
Db          1 MVAGSYRRYEFVRNRDDSDDGLKPLKDLNGQICQICGDDVGLTKTGNVFVACNECGFPLC 60

Qy         67 RDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAES 126
              : |||||::|:| ||||| |::|  |  ||  ||:|| |:|::|||:
Db         61 QSCYEYERKDGSQCCPQCKARFRRHNGSPRVEVDEKEDDVNDIENEFD------------ 108

Qy        127 MLYGHMSYGRGGDPNGAP---QAFQLNPN------VPLLTNGQMVD-DIP-PEQHALVPS 175
                  | :| :    |   : |  :         | |||:|  |  :|| |:::| :
Db        109 -------YTQGNNKARLPHRAEEFSSSSRHEESLPVSLLTHGHPVSGEIPTPDRNATLS- 160

Qy        176 FMGGGGKRIHPLPYADPSLP--------VQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ 228
                      |  || ||       :  | :|||||| :||  :| ||:|:: || :|
Db        161 ------------PCIDPQLPGIYQLLLLPVRILDPSKDLNSYGLVNVDWKKRIQGWKLKQ 208

Qy        229 ER--MHQTGN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIII 277
              ::  :| ||          :| | :||  :| ::|:||  :|| :  ||::: |||::|:
Db        209 DKNMIHMTGKYHEGKGGEFEGTGSNGD--ELQMVDDARLPMSRVVHFPSARMTPYRIVIV 266

Qy        278 IRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLS 337
              :||::|| | |||  ||| ||:|||| ||||||||| ||:||||||:|| |||:||||:
Db        267 LRLIILGVFLHYRTTHPVKDAYALWLTSVICEIWFAFSWLLDQFPKWYPINRETFLDRLA 326

Qy        338 LRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAM 397
              ||:|::|:|||||:| |||||||:||||||| |||||||:|||||||:|||||||:||
Db        327 LRYDRDGEPSQLAPVDVFVSTVDPMKEPPLVTANTVLSILAVDYPVDKVACYVSDDGSAM 386

Qy        398 LTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREY 457
              |||||||||:||:|||||||::|||||||:|| |||||||||:  :||:|||||||||
Db        387 LTFEALSETAEFSKKWVPFCKKFNIEPRAPEFYFSQKIDYLKDKIQPSFVKERRAMKREY 446

Qy        458 EEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELP 517
              |||||||| |||||||:||:||||:||| ||||| |||||||||||| ||||| :|||||
Db        447 EEFKVRINILVAKAQKIPEDGWTMEDGTSWPGNNPRDHPGMIQVFLGHSGGLDTDGNELP 506

Qy        518 RLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFM 577
              ||:|||||||||: |||||||||||:|||||||| ||||:||||| ||||||||||||||
Db        507 RLIYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNNSKAIKEAMCFM 566

Qy        578 MDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQ 637
              ||| :||| |||||||||||||| |||||||:|||||||:|||||| | ||
Db        567 MDPAIGKKCCYVQFPQRFDGIDLHDRYANRNTVFFDINLKGLDGIQGPVYVGTGCCFNRQ 626

Qy        638 ALYGYDAPKTKK--PPSRTCNCWPKWCFCCCCFGNRKQKKTTK-PKTEKKKLLFFKKEEN 694
              ||||||    |::   |:              |||:||: |: | |  |  : :  |: ::
Db        627 ALYGYDPVLTEEDLEPN---------IIVKSCFGSRKKGKSRKIPNYEDNRSI--KRSDS 675

Qy        695 QSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASL 754
              | : : :|||   | |:| : :|:|::|||:|||| ||: :| :| ||   : :| :|
Db        676 NVPLFNMEDIDEDVEGYEDEMSLLVSQKRLEKRFGQSPVFIAATFMEQGGLPSTTNPLTL 735

Qy        755 LKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSA 814
              |||||||||||| ||||||||||||||||||||||||   || ||||:| | ||||||
Db        736 LKEAIHVISCGYEAKTDWGKEIGWIYGSVTEDILTGFKMHARGWISIYCVPSRPAFKGSA 795

Qy        815 PLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLA 874
              |:|||||:|||||||||||  | |||:|||| | || ||| :|||:|||| |||||||
Db        796 PINLSDRLNQVLRWALGSIEILLSRHCPIWYGYNGRLKLLERIAYINTIVYPITSIPLLA 855

Qy        875 YCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWV 934
              || ||| ||:|  || ||::|:||| || ||  |:|::||||::|| | ::||||||||||
Db        856 YCMLPAFCLITNTFIIPEISNLASLCFMLLFASIYASAILELKWSDVALEDWWRNEQFWV 915

Qy        935 IGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLL 993
              ||| |:||||||||||||| ||:||:||||||  | | :|:||| ||||:|||||||:||:

Db        916 IGGTSAHLFAVFQGLLKVFAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTILLV 975

Qy        994 NFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSI 1053
              | :|:||||| |||:||:||||| ||| |||||: |||||||||:||||||||||||||
Db        976 NLVGIVAGVSYAINSGYQSWGPLMGKLLFAFWVVAHLYPFLKGLLGRQNRTPTIVIVWSA 1035

Qy       1054 LLASIFSLLWVRIDPFLA 1071
              ||||||||||||||:||::
Db       1036 LLASIFSLLWVRINPFVS 1053


RESULT 8
T10797
cellulose synthase (EC 2.4.1.-) catalytic chain celA1 - upland cotton
C;Species: Gossypium hirsutum (upland cotton)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10797
R;Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A;Title: Higher plants contain homologs of the bacterial celA genes encoding the catal
A;Reference number: Z17152; MUID:97057296; PMID:8901635
A;Accession: T10797
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-974 <PEA>
A;Cross-references: EMBL:U58283; NID:g1706955; PIDN:AAB37766.1; PID:g1706956
A;Experimental source: strain Acala SJ-2; fiber
C;Genetics:
A;Gene: celA1
C;Function:
A;Description: involved in the synthesis of cellulose
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             58.6%;  Score 3469.5;  DB 2;  Length 974;
  Best Local Similarity   59.8%;  Pred. No. 9.6e-258;
  Matches  638;  Conservative  152;  Mismatches  168;  Indels  109;  Gaps   13;

Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKR-- 90
              |    ||  ||: |||   |:|||||:|| ||:|: |:||: :|| : | :| : |
Db          3 ESGVPVCHTCGEHVGLNVNGEPFVACHECNFPICKSCFEYDLKEGRKACLRCGSPYDENL 62

Qy         91 LKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLN 150
              |   :: |||                          :| :  |::  :
Db         63 LDDVEKATGD-----------------------QSTMAAHLNKSQD------------- 85

Qy        151 PNVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAY 210
                                          :|   : |  :   |  :             |
Db         86 -------------------------VGIHARHISSVSTLDSEM--------------AE 105

Qy        211 GYGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLP----LMD----EARQQLSRKI 262
                |:  || |:|:||::: :  :           :|::|    : |     :| | ||  |
Db        106 DNGNSIWKNRVESWKEKKNKKKKPATT---KVEREAEIPPEQQMEDKPAPDASQPLSTII 162

Qy        263 PLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFP 322
              |:| |:: ||| :||:||::|| |||||| :||: || ||| ||||||||| ||:||||
Db        163 PIPKSRLAPYRTVIIMRLIILGLFFHYRVTNPVDSAFGLWLTSVICEIWFAFSWVLDQFP 222

Qy        323 KWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYP 382
              ||:|: ||||:|||| |:::||:| :|| :|||||||||||||:| |||||||::|||
Db        223 KWYPVNRETYIDRLSARYEREGEPDELAAVDFFVSTVDPLKEPPLITANTVLSILALDYP 282

Qy        383 VDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKV 442
              |||||||:|||||||||||:| ||::||:|||||||:::|||||||:|| ||||||||||
Db        283 VDKVSCYISDDGAAMLTFESLVETADFARKWVPFCKKFSIEPRAPEFYFSQKIDYLKDKV 342

Qy        443 AANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVF 502
              :||:|||||||:|||:|:|||||||||| |:|||||||| |||| ||||||||||
Db        343 QPSFVKERRAMKRDYEEYKIRINALVAKAQKTPDEGWTMQDGTSWPGNNPRDHPGMIQVF 402

Qy        503 LGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDH 562
              || ||   | |||||||||||||||||| ||||||| |||||||||||||::|||||||
Db        403 LGYSGARDIEGNELPRLVYVSREKRPGYQHHKKAGAENALVRVSAVLTNAPFILNLDCDH 462
```

```
Qy        563 YINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGI 622
Db        463 YVNNSKAVREAMCFLMDPQVGRDVCYVQFPQRFDGIDRSDRYANRNTVFFDVNMKGLDGI 522

Qy        623 QGPIYVGTGCVFRRQALYGYDAPKTKK-PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKT 681
Db        523 QGPVYVGTGCVFNRQALYGYGPPSMPSFPKSSSSSC------SCCCPGKK------EPKD 570

Qy        682 EKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLE 741
Db        571 PSELYRDAKREELDAAIFNLREIDNY---DEYERSMLISQTSFEKTFGLSSVFIESTLME 627

Qy        742 NGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSI 801
Db        628 NGGVAESANPSTLIKEAIHVISCGYEEKTAWGKEIGWIYGSVTEDILTGFKMHCRGWRSI 687

Qy        802 YCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG-LKFLERFSYI 860
Db        688 YCMPLRPAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPLWYGFGGGRLKWLQRLAYI 747

Qy        861 NSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSG 920
Db        748 NTIVYPFTSLPLIAYCSLPAICLLTGKFIIPTLSNLASVLFLGLFLSIIVTAVLELRWSG 807

Qy        921 VGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKW 980
Db        808 VSIEDLWRNEQFWVIGGVSAHLFAVFQGFLKMLAGIDTNFTVTAKAADDADFGELYIVKW 867

Qy        981 TTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGR 1040
Db        868 TTLLIPPTTLLIVNMVGVVAGFSDALNKGYEAWGPLFGKVFFSFWVILHLYPFLKGLMGR 927

Qy       1041 QNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE--CGLDC 1085
Db        928 QNRTPTIVVLWSVLLASVFSLVWVRINPFVSTADSTTVSQSCISIDC 974
```

RESULT 9
T04870
cellulose synthase (EC 2.4.1.-) catalytic chain F28A21.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04870
R;Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15387
A;Accession: T04870
A;Molecule type: DNA
A;Residues: 1-958 <BEV>
A;Cross-references: EMBL:AL035526
A;Experimental source: cultivar Columbia; BAC clone F28A21
C;Genetics:
A;Map position: 4
A;Introns: 60/1; 76/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791
A;Note: F28A21.190
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match 56.3%; Score 3329.5; DB 2; Length 958;
  Best Local Similarity 58.0%; Pred. No. 5.3e-247;
  Matches 622; Conservative 150; Mismatches 165; Indels 135; Gaps 14;

```
Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLK 92
Db          3 ESRSPICNTCGEEIGVKSNGEFFVACHECSFPICKACLEYEFKEGRRICLRCGNPY---- 58

Qy         93 GCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPN 152
Db         59 ---------DENVFDDVETKTS----KTQSIDSGIHARHIS------------------- 86

Qy        153 VPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGY 212
Db         87 ----TVSTIDSELNDE-------------------------------------------Y 99
```

```
Qy        213 GSVAWKERMENWKQRQERMHQTGN---DGGGDDGDDADLPLMDE----------ARQQL 258
Db        100 GNPIWKNRVESWKDKKDKKSKKKKKDPKATKAEQHEAQIPTQQHMEDTPPNTESGATDVL 159

Qy        259 SRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWIL 318
Db        160 SVVIPIPRTKITSYRIVIIMRLIILALFFNYRITHPVDSAYGLWLTSVICEIWFAVSWVL 219

Qy        319 DQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILS 378
Db        220 DQFPKWSPINRETYIDRLSARFEREGEQSQLAAVDFFVSTVDPLKEPPLITANTVLSILA 279

Qy        379 VDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYL 438
Db        280 LDYPVDKVSCYVSDDGAAMLSFESLVETADFARKWVPFCKKYSIEPRAPEFYFSLKIDYL 339

Qy        439 KDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGM 498
Db        340 RDKVQPSFVKERRAMKRDYEEFKIRMNALVAKAQKTPEEGWTMQDGTSWPGNNTRDHPGM 399

Qy        499 IQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNL 558
Db        400 IQVFLGYSGARDIEGNELPRLVYVSREKRPGYQHHKKAGAENALVRVSAVLTNAPFILNL 459

Qy        559 DCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKG 618
Db        460 DCDHYVNNSKAVREAMCFLMDPVVGQDVCFVQFPQRFDGIDKSDRYANRNIVFFDVNMRG 519

Qy        619 LDGIQGPIYVGTGCVFRRQALYGYDAPKTKK--PPSRTCNCWPKWCFCCCCFGNRKQKKT 676
Db        520 LDGIQGPVYVGTGTVFRRQALYGYSPPSKPRILPQSSSSS--------CCCLTKKKQ--- 568

Qy        677 TKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVT 736
Db        569 --PQDPSEIYKDAKREELDAAIFNLGDLDNY---DEYDRSMLISQTSFEKTFGLSTVFIE 623

Qy        737 STLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCH 796
Db        624 STLMENGGVPDSVNPSTLIKEAIHVISCGYEEKTEWGKEIGWIYGSITEDILTGFKMHCR 683

Qy        797 GWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG-LKFLE 855
Db        684 GWRSIYCMPLRPAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPLWYGCSGGRLKLLQ 743

Qy        856 RFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILE 915
Db        744 RLAYINTIVYPPTSLPLVAYCTLPAICLLTGKFIIPTLSNLASMLFL------------- 790

Qy        916 MRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSEL 975
Db        791 ----GVSIEDLWRNEQFWVIGGVSAHLFAVFQGFLKMLAGLDTNFTVTSKTADDLEFGEL 846

Qy        976 YTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLK 1035
Db        847 YIVKWTTLLIPPTSLLIINLVGVVAGFSDALNKGYEAWGPLFGKVFFAFWVILHLYPFLK 906

Qy       1036 GLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLE-ECGL-DC 1085
Db        907 GLMGRQNRTPTIVILWSILLASVFSLVWVRINPFVSKTDTTSLSLNCLLIDC 958
```

RESULT 10
T10800
cellulose synthase (EC 2.4.1.-) catalytic chain celA2 - upland cotton (fragment)
C;Species: Gossypium hirsutum (upland cotton)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10800
R;Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A;Title: Higher plants contain homologs of the bacterial celA genes encoding the catal
A;Reference number: Z17152; MUID:97057296; PMID:8901635
A;Accession: T10800
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA
A;Residues: 1-685 <PEA>
A;Cross-references: EMBL:U58284; NID:g1706957; PIDN:AAB37767.1; PID:g1706958
A;Experimental source: strain Acala SJ-2; fiber
C;Genetics:
A;Gene: celA2
C;Function:
A;Description: involved in synthesis of cellulose
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.1%; Score 2907; DB 2; Length 685;
Best Local Similarity 75.0%; Pred. No. 8.9e-215;
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

```
Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470
              ::||||||::|:||||||:| :|||||||||  :||:||||||||||||||||||||||
Db          1 RRWVPFCKKHNVEPRAPEFYFNEKIDYLKDKVHPSFVKERRAMKREYEEFKVRINALVAK 60

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530
              ||| ||||| |||||||||| |||||||||:|| :| || :| ||||||||||||||||
Db         61 AQKKPEEGWVMQDGTPWPGNNTRDHPGMIQVYLGSAGALDVDGKELPRLVYVSREKRPGY 120

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590
              ||||||| |||||||||||||::|||||||||||||::|||||:|||  |||:||||
Db        121 QHHKKAGAENALVRVSAVLTNAPFILNLDCDHYINNSKAMREAMCFLMDPQFGKKLCYVQ 180

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
              |||||||||||||||||||||||| ||||:|||:|||||||| |||||||| | ::|
Db        181 FPQRFDGIDRHDRYANRNVVFFDINMLGLDGLQGPVYVGTGCVFNRQALYGYDPPVSEKR 240

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKK----KLLFFKK---------EENQSP 697
              |  ||:||| || |||| |:||:   :| | |||     ||: ||          ::  :|
Db        241 PKMTCDCWPSWC-CCCCGGSRKK---SKKKGEKKGLLGGLLYGKKKKMMGKNYVKKGSAP 296

Qy        698 AYALGEIDEAAPGAEN-EKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLK 756
              : | ||:|   | |  ||: :::|:  ||:|||| ||: |||:|||| :  :  ||:|
Db        297 VFDLEEIEEGLEGYEELEKSTLMSQKNFEKRFGQSPVFIASTLMENGGLPEGTNSTSLIK 356

Qy        757 EAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPL 816
              |||||||||||:||:|||||||||||||||||||||||| ||:|:||:||| |||||||:
Db        357 EAIHVISCGYEEKTEWGKEIGWIYGSVTEDILTGFKMHCRGWKSVYCVPKRPAFKGSAPI 416

Qy        817 NLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYC 876
              |||||||||||||||:||| | |||||||||| ||:||| :|||:||||:|||||||||
Db        417 NLSDRLHQVLRWALGSVEIFLSRHCPLWYGYGGKLKWLERLAYINTIVYPFTSIPLLAYC 476

Qy        877 TLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIG 936
              |:||:|||||||| | |:|: |:||::||: | || :||:|||| | ||||||||||||
Db        477 TIPAVCLLTGKFIIPTLSNLTSVWFLALFLSIIATGVLELRWSGVSIQDWWRNEQFWVIG 536

Qy        937 GVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFI 996
              |||:||||||||||||:|||||:||||:|  || || ||| ||||||||||||::|| :
Db        537 GVSAHLFAVFQGLLKVLAGVDTNFTVTAKAADDTEFGELYLFKWTTLLIPPTTLIILNMV 596

Qy        997 GVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLA 1056
              |||||||:|||||| ||||||||||||||:|||||||||:||||||||||::||||||
Db        597 GVVAGVSDAINNGYGSWGPLFGKLFFAFWVILHLYPFLKGLMGRQNRTPTIVVLWSILLA 656

Qy       1057 SIFSLLWVRIDPFLAKDDGPLLEECGLDC 1085
              |||||:|||||||| |  ||:|::||::|
Db        657 SIFSLVWVRIDPFLPKQTGPVLKQCGVEC 685
```

RESULT 11
D86157
hypothetical protein F22D16.26 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: D86157
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86157
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1181 <STO>
A;Cross-references: GB:AE005172; NID:g6056428; PIDN:AAF02892.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 39.4%; Score 2330.5; DB 2; Length 1181;
Best Local Similarity 43.0%; Pred. No. 3.5e-170;
Matches 489; Conservative 166; Mismatches 274; Indels 207; Gaps 26;

```
Qy         31 PREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKR 90
              | :::||:| : | |  :  |      | || | :|||||      |  ||| ||  |:
Db        154 PEKKSGQICWLKGCDEKVVHG-----RC-ECGFRICRDCYFDCITSGGGNCPGCKEPYRD 207

Qy         91 LKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLN 150
              :        ::|||    |            : || |   :|
Db        208 INDDPETEEEDEEDEAKPL-----------PQMGESKLDKRLS---------------- 239

Qy        151 PNVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDL--- 207
                                :| ||                    |    | :: |
Db        240 --------------------VVKSFKAQN---------------QAGDFDHTRWLFET 262

Qy        208 -AAYGYGSVAWKERMENWKQRQERMHQTG---NDGGGDDGDDADLPLMDEARQQLSRKIP 263
                 ||||:  |                : |     ||| :| :      : ::: |:||:
Db        263 KGTYGYGNAVWP-------------KDGYGIGSGGGGNGYETPPEFGERSKRPLTRKVS 308

Qy        264 LPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPK 323
              : :: |:|||::| :||| || |  :|| ||  :|  || :|  ||:|||:|||||| ||
Db        309 VSAAIISPYRLLIALRLVALGLFLTWRVRHPNREAMWLWGMSTTCELWFALSWLLDQLPK 368

Qy        324 WFPIERETYLDRLSLRFDK------EGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSIL 377
                |: | |  |  ||:       :|: | |  || |||| || ||||||| ||:||||
Db        369 LCPVNRLTDLGVLKERFESPNLRNPKGR-SDLPGIDVFVSTADPEKEPPLVTANTILSIL 427

Qy        378 SVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDY 437
              :|||||:|::||:|||| |:||||||::|: ||  |||||:::||||| || || || ::
Db        428 AVDYPVEKLACYLSDDGGALLTFEALAQTASFASTWVPFCRKHNIEPRNPEAYFGQKRNF 487

Qy        438 LKDKVAANFVRERRAMKREYEEFKVRINAL---------------VAKAQK--------- 473
              ||:||  :|||||| :||||:|||||||:|                  :|:|
Db        488 LKNKVRLDFVRERRRVKREYDEFKVRINSLPEAIRRRSDAYNVHEELRAKKKQMEMMMGN 547

Qy        474 -------VPEEGWTMQDGTPWPG--------NNVRDHPGMIQVFLGQSGGLDCEGNE--- 515
                    ||:  | | ||: |||          |:  || |:||  |          | |
Db        548 NPQETVIVPKATW-MSDGSHWPGTWSSGETDNSRGDHAGIIQAMLAPPNAEPVYGAEADA 606

Qy        516 ------------LPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHY 563
                         || ||||||||||||:|:|||||||||| ||:::| |::||||||||
Db        607 ENLIDTTDVDIRLPMLVYVSREKRPGYDHNKKAGAMNALVRTSAIMSNGPFILNLDCDHY 666

Qy        564 INNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQ 623
              | || |::| ||||:|   | ::|||||||||:||| :|||| | ||||::|: |||:|
Db        667 IYNSMALREGMCFMLD-RGGDRICYVQFPQRFEGIDPNDRYANHNTVFFDVSMRALDGLQ 725

Qy        624 GPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEK 683
              ||:||||||:||| ||||:  |:  :          |       | || | : :    :
Db        726 GPMYVGTGCIFRRTALYGFSPPRATEHHG--------W------LGRRKVKISLR----R 767

Qy        684 KKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTS-TLLEN 742
              | :  | :|   |    || :|     |    | :    | |:|| |: || |  : |
Db        768 PKAMMKKDDEVSLPIN--GEYNE-----EENDDGDIESLLLPKRFGNSNSFVASIPVAEY 820

Qy        743 GGTL--------KSASPASLL-----------KEAIHVISCGYEDKTDWGKEIGWIYGS 782
              | |        |:: ||  |            ||| |||| |||||:||| :||||||
```

```
Db        821 QGRLIQDLQGKGKNSRPAGSLAVPREPLDAATVAEAISVISCFYEDKTEWGKRVGWIYGS 880

Qy        783 VTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCP 842

Db        881 VTEDVVTGYRMHNRGWRSIYCVTKRDAFRGTAPINLTDRLHQVLRWATGSVEIFFSRNNA 940

Qy        843 LWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFM 902

Db        941 IF--ATRRMKFLQRVAYFNVGMYPFTSLFLIVYCILPAISLFSGQFIVQSLDITFLIYLL 998

Qy        903 SLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTV 962

Db        999 SITLTLCMLSLLEIKWSGITLHEWWRNEQFWVIGGTSAHPAAVLQGLLKVIAGVDISFTL 1058

Qy        963 TSKGG----DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFG 1018

Db       1059 TSKSSAPEDGDDEFADLYVVKWSFLMVPPLTIMMVNMIAIAVGLARTLYSPFPQWSKLVG 1118

Qy       1019 KLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDD 1074

Db       1119 GVFFSFWVLCHLYPFAKGLMGRRGRVPTIVFVWSGLLSIIVSLLWVYINPPSGKQD 1174
```

RESULT 12
T51546
cellulose synthase catalytic subunit-like protein - Arabidopsis thaliana
N;Alternate names: protein F2K13_60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C;Accession: T51546
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1145 <SAT>
A;Cross-references: EMBL:AL391141
A;Experimental source: cultivar Columbia; BAC clone F2K13
C;Genetics:
A;Map position: 5
A;Introns: 297/2; 566/3
A;Note: F2K13_60

  Query Match 39.1%; Score 2311; DB 2; Length 1145;
  Best Local Similarity 42.6%; Pred. No. 1e-168;
  Matches 493; Conservative 170; Mismatches 271; Indels 222; Gaps 33;

```
Qy         11 SHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPG--GDPFVACNECAFPVCRD 68

Db        100 SNTRAHLMHKVIETEPNHPQMAGSKGSSCAIPGCDAKVMSDERGQDLLPC-ECDFKICRD 158

Qy         69 CYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESML 128

Db        159 CFIDAVKTGGGICPGCKEPYKN----THLTDQVDENG---------------------- 191

Qy        129 YGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGGG----KRI 184

Db        192 ----------------------------------------QQRPMLP---GGGGSKMERRL 209

Qy        185 HPLPYADPSLPVQPRSMDPS------KDLAAYGYGSVAWKERMENWKQRQERMHQTGNDG 238

Db        210 SMVKSTNKSALMRSQTGDFDHNRWLFETTGTYGYGNAFWT-------------KDGDFG 255

Qy        239 GGDDGD-DAD------LPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRV 291

Db        256 SGKDGDGDGDGMGMEAQDLMSRPWRPLTRKLKIPAGVISPYRLLIFIRIVVLALFLTWRV 315

Qy        292 MHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFD--KEGQP--- 346

Db        316 KHQNPDAVWLWGMSVVCELWFALSWLLDQLPKLCPINRATDLQVLKEKFETPTASNPTGK 375

Qy        347 SQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSET 406

Db        376 SDLPGFDVFVSTADPEKEPPLVTANTILSILAAEYPVEKLSCYVSDDGGALLTFEAMAEA 435

Qy        407 SEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINA 466

Db        436 ASFANIWVPFCRKHAIEPRNPDSYFSLKRDPYKNKVKSDFVKDRRRVKREFDEFKVRVNS 495

Qy        467 L------------------VAKAQ-----------KVPEEGWTMQDGTPWPGNNVR--- 493

Db        496 LPDSIRRRSDAYHAREEIKAMKMQRQNRDDEPMEPVKIPKATW-MADGTHWPGTWLTSAS 554

Qy        494 -----DHPGMIQVFL---------GQSGG-LDCEGNE--LPRLVYVSREKRPGYNHHKKA 536

Db        555 DHAKGDHAGIIQVMLKPPSDEPLHGVSEGFLDLTDVDIRLPLLVYVSREKRPGYDHNKKA 614

Qy        537 GAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFD 596

Db        615 GAMNALVRASAIMSNGPFILNLDCDHYIYNSEALREGMCFMMD-RGGDRLCYVQFPQRFE 673

Qy        597 GIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK-PPSRTC 655

Db        674 GIDPSDRYANHNTVFFDVNMRALDGLMGPVYVGTGCLFRRIALYGFNPPRSKDFSPS--- 730

Qy        656 NCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEK 715

Db        731 -CW------SCCF----------PRSKKKNI----PEENR--ALRMSDYDDEE------- 760

Qy        716 AGIVNQQKLEKKFGQSSVFVTS--------------TLLENG---GTL---KSASPASLL 755

Db        761 ---MNLSLVPKKFGNSTFLIDSIPVAEFQGRPLADHPAVKNGRPPGALTIPRELLDASTV 817

Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815

Db        818 AEAIAVISCWYEDKTEWGSRIGWIYGSVTEDVVTGYRMHNRGWKSVYCVTKRDAFRGTAP 877

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875

Db        878 INLTDRLHQVLRWATGSVEIFFSRNNALL--ASSKMKILQRIAYLNVGIYPFTSIFLIVY 935

Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935

Db        936 CFLPALSLFSGQFIVQTLNVTFLVYLLIISITLCLLALLEIKWSGISLEEWWRNEQFWLI 995

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSK-GGD--DEEFSELYTFKWTTLLIPPTTLLL 992

Db        996 GGTSAHLAAVLQGLLKVVAGVEISFTLTSKSGGDDIDDEFADLYMVKWTSLMIPPITIIM 1055

Qy        993 LNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWS 1052

Db       1056 VNLIAIAVGFSRTIYSVVPQWSKLIGGVFFSFWVLAHLYPFAKGLMGRRGRTPTIVYVWS 1115

Qy       1053 ILLASIFSLLWVRIDP 1068

Db       1116 GLVAITISLLWVAINP 1131
```

RESULT 13
T05646
hypothetical protein F20D10.310 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T05646
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05646
A;Molecule type: DNA
A;Residues: 1-1111 <BEV>
A;Cross-references: EMBL:AL035538
A;Experimental source: cultivar Columbia; BAC clone F20D10
C;Genetics:
A;Map position: 4
A;Introns: 139/2; 675/3

A;Note: F20D10.310

Query Match 38.1%; Score 2256; DB 2; Length 1111;
Best Local Similarity 41.7%; Pred. No. 1.7e-164;
Matches 472; Conservative 163; Mismatches 260; Indels 236; Gaps 27;

```
Qy        36 GQVCQI--CGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKG 93
             |  | :  |  :|     |   : | || | :||||:   ::| |  || || :||
Db       112 GSSCAMPACDGNVMKDERGKDVMPC-ECRFKICRDCFMDAQKE-TGLCPGCKEQYK---- 165

Qy        94 CQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQL--NP 151
                || ::|  |                              |  |  |  ||   |   |
Db       166 ----IGDLDDDTPD-------------------------YSSGALPLPAPGKDQRGNNN 195

Qy       152 NVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDL---- 207
             |: ::   |                          |:                  | :: |
Db       196 NMSMMKRNQ------------------NGE-----------------FDHNRWLFETQ 218

Qy       208 AAYGYGSVAWKERMENWKQRQERMHQTGNDGGGDDGDD----ADLPLMDEARQQLSRKIP 263
               ||||:  |          |: |:       ||| |:      :   |:  : |||:||
Db       219 GTYGYGNAYWP---------QDEMY-------GDDMDEGMRGGMVETADKPWRPLSRRIP 262

Qy       264 LPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPK 323
             :|:: |:|||::|:|| ||| ||  :|: :|  ||  |||:|:|||:||  ||||| ||
Db       263 IPAAIISPYRLLIVIRFVVLCFFLTWRIRNPNEDAIWLWLMSIICELWFGFSWILDQIPK 322

Qy       324 WFPIERETYLDRLSLRFD--KEGQP---SQLAPIDFFVSTVDPLKEPPLVTTNTVLSILS 378
               || | |: |  :||      |   | |  || |||| || ||||||| ||:||||:
Db       323 LCPINRSTDLEVLRDKFDMPSPSNPTGRSDLPGIDLFVSTADPEKEPPLVTANTILSILA 382

Qy       379 VDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYL 438
             |||||:|||||:|||| |:|:|||::| : ||  |||||:::|||| |: ||  |||
Db       383 VDYPVEKVSCYLSDDGGALLSFEAMAEAASFADLWVPFCRKHNIEPRNPDSYFSLKIDPT 442

Qy       439 KDKVAANFVRERRAMKREYEEFKVRINAL----------------------------VA 469
             |:|   :||::|| :||||:|||||| |
Db       443 KNKSRIDFVKDRRKIKREYDEFKVRINGLPDSIRRRSDAFNAREEMKALKQMRESGGDPT 502

Qy       470 KAQKVPEEGWTMQDGTPWPG--------NNVRDHPGMIQVFLGQSGGLDCEGN------- 514
             :  |||:  | | ||| |||        ::  || |::|| |          ||
Db       503 EPVKVPKATW-MADGTHWPGTWAASTREHSKGDHAGILQVMLKPPSSDPLIGNSDDKVID 561

Qy       515 ------ELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSK 568
                   ||  |||||||||||:|:|||||||||| ||:|:| |::|||||||| | |
Db       562 FSDTDTRLPMFVYVSREKRPGYDHNKKAGAMNALVRASAILSNGPFILNLDCDHYIYNCK 621

Qy       569 AIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYV 628
             |::| ||||||   |: :||:|||||:|||  ||||| | |||| ||: |||:|||:||
Db       622 AVREGMCFMMD-RGGEDICYIQFPQRFEGIDPSDRYANNNTVFFDGNMRALDGVQGPVYV 680

Qy       629 GTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLF 688
             ||| :||| ||||:| |   |
Db       681 GTGTMFRRFALYGFDPPNPDK-----------------------------------L 702

Qy       689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTS--------TLL 740
              :|:|::: |    : |              ::  :| |:|| |::   |          |
Db       703 LEKKESETEALTTSDFDPD-----------LDVTQLPKRFGNSTLLAESIPIAEFQGRPL 751

Qy       741 ENGGTLKSASP------------ASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDIL 788
              :   :|   |            |: : |:: |||| |||||:||  :||||||||||::
Db       752 ADHPAVKYGRPPGALRVPRDPLDATTVAESVSVISCWYEDKTEWGDRVGWIYGSVTEDVV 811

Qy       789 TGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYG 848
             ||:|||  ||||:||| || :|:||||:||:|||||||||| ||:|||||  :  :
Db       812 TGYRMHNRGWRSVYCITKRDSFRGSAPINLTDRLHQVLRWATGSVEIFFSRNNAIL--AS 869

Qy       849 GGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICI 908
              ||||:| :|:|  :||:||: |: || |||  | :|:||   |:    :: : :  ||:
Db       870 KRLKFLQRLAYLNVGIYPFTSLFLILYCFLPAFSLFSGQFIVRTLSISFLVYLLMITICL 929

Qy       909 FATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSK-GG 967
              ::||::|||:|:::|||||:|:| | ||||:|| ||:||||||:: |||:|:| ||
Db       930 IGLAVLEVKWSGIGLEEWWRNEQWWLISGTSSHLYAVVQGVLKVIAGIEISFTLTTKSGG 989
```

```
Qy       968 DDEE--FSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFW 1025
             || |  :::|| ||::|:|||  : ::| | :|      |       |  | |  ||:||
Db       990 DDNEDIYADLYIVKWSSLMIPPIVIAMVNIIAIVVAFIRTIYQAVPQWSKLIGGAFFSFW 1049

Qy      1026 VIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGP 1076
             |: ||||| |||:||: :||||| ||: |:|   ||||  |:|    : ||
Db      1050 VLAHLYPFAKGLMGRRGKTPTIVFVWAGLIAITISLLWTAINP----NTGP 1096
```

RESULT 14
D84741
probable cellulose synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: D84741
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84741
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1036 <STO>
A;Cross-references: GB:AE002093; NID:g2924781; PIDN:AAC04910.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g33100
A;Map position: 2

Query Match 36.6%; Score 2166.5; DB 2; Length 1036;
Best Local Similarity 44.8%; Pred. No. 1.1e-157;
Matches 447; Conservative 160; Mismatches 222; Indels 169; Gaps 28;

```
Qy       178 GGGG----------KRIHPLPYADPSLPVQPRSMD--------PSKDLAAYGYGSVAWKE 219
             ||||           :|:  :   : |: :: :: |           ||    || |:  | |
Db        89 GGGGDGPKMGNKLERRLSVMKSNNKSMLLRSQTGDFDHNRWLFESK--GKYGIGNAFWSE 146

Qy       220 RMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIR 279
               : :             |||    |      :|:  : |:||: :|:  ::|||::|:||
Db       147 EDDTY------------DGGVSKSD-----FLDKPWKPLTRKVQIPAKILSPYRLLIVIR 189

Qy       280 LVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLR 339
             ||:: ||  :|: :|  ||  || :|::|||||| |||||  ||  || | | |  |
Db       190 LVIVFFFLWWRITNPNEDAMWLWGLSIVCEIWFAFSWILDILPKLNPINRATDLAAL--- 246

Qy       340 FDKEGQPSQLAP--------IDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVS 391
              ||  |||    |         :| |||| || ||||||| ||:||||:||||::|:| |:|
Db       247 HDKFEQPSPSNPTGRSDLPGVDVFVSTADPEKEPPLVTANTLLSILAVDYPIEKLSAYIS 306

Qy       392 DDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERR 451
             ||| |:|||||::|   ||: |||||::::|||| |: ||  | |  |:|   :||::||
Db       307 DDGGAILTFEAMAEAVRFAEYWVPFCRKHDIEPRNPDSYFSIKKDPTKNKKRQDFVKDRR 366

Qy       452 AMKREYEEFKVRINALVAKAQK----------------------VPEEG-------WT 480
              :||||:|||||| |  : :|                      :| :|       |
Db       367 WIKREYDEFKVRINGLPEQIKKRAEQFNMREELKEKRIAREKNGGVLPPDGVEVVKATW- 425

Qy       481 MQDGTPWPG--------NNVRDHPGMIQVF---------LG--QSGGLDCEGNEL--PRL 519
             | ||| |||          ::  || |::|:           :|    | ||  | ::  |
Db       426 MADGTHWPGTWFEPKPDHSKGDHAGILQIMSKVPDLEPVMGGPNEGALDFTGIDIRVPMF 485

Qy       520 VYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMD 579
             |||||||||::|:|||||| :|| ||:|:|  ::|||||||| |||||| ||||||
Db       486 AYVSREKRPGFDHNKKAGAMNGMVRASAILSNGAFILNLDCDHYIYNSKAIKEGMCFMMD 545

Qy       580 PLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQAL 639
              | ::||:||||||:|||  |||||  | |||| ||: |||:|||:||||||:||| ||
Db       546 -RGGDRICYIQFPQRFEGIDPSDRYANHNTVFFDGNMRALDGLQGPVYVGTGCMFRRYAL 604

Qy       640 YGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAY 699
             ||:: |:  :                 ||            ||   :  : :   |
Db       605 YGFDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAY 664
```

```
Db        605 YGFNPPRANEYSG--------------VFGQ-----------EKAPAMHVRTQSQASQTS 639

Qy        700 ALGEID-EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTL--------------LENG 743
                  ::: :  |  ::   |       | |||| |::| | |:                ::||
Db        640 QASDLESDTQPLNDDPDLG------LPKKFGNSTMF-TDTIPVAEYQGRPLADHMSVKNG 692

Qy        744 ---GTLKSASP---ASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHG 797
                |  |    |   |  : ||| |||| ||| |:||  ||||||||||||::||::||  |
Db        693 RPPGALLLPRPPLDAPTVAEAIAVISCWYEDNTEWGDRIGWIYGSVTEDVVTGYRMHNRG 752

Qy        798 WRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERF 857
              |||:||| || ||:|:||:||:|||||||||| ||:|||| :  ::      ||||:|
Db        753 WRSVYCITKRDAFRGTAPINLTDRLHQVLRWATGSVEIFFSKNNAMF--ATRRLKFLQRV 810

Qy        858 SYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFAT----SI 913
              :|:|  :||:||| |: || |||:|| :||||   |:    : |:|  :||  |    |:
Db        811 AYLNVGIYPFTSIFLVVYCFLPALCLFSGKFIVQSLD----IHFLSYLLCITVTLTLISL 866

Qy        914 LEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSK-GGDDEE- 971
              ||:|||:|:::|||||||:||| |:|| || |||||||:: |||:|||  |:||:
Db        867 LEVKWSGIGLEEWWRNEQFWLIGGTSAHLAAVVQGLLKVIAGIEISFTLTSKASGEDEDD 926

Qy        972 -FSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHL 1030
               |:||  ||| | | | |::::| : :| | |  | :    || | | :||: ||: |:
Db        927 IFADLYIVKWTGLFIMPLTIIIVNLVAIVIGASRTIYSVIPQWGKLMGGIFFSLWVLTHM 986

Qy       1031 YPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
              ||| |||:||: : |||| ||| |::   ||||: | |
Db        987 YPFAKGLMGRRGKVPTIVYVWSGLVSITVSLLWITISP 1024
```

RESULT 15
T02209
cellulose synthase (EC 2.4.1.-) catalytic chain - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 11-Jun-1999
C;Accession: T02209
R;Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H
Science 279, 717-720, 1998
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Reference number: Z13745; MUID:98111412; PMID:9445479
A;Accession: T02209
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-583 <ARI>
A;Cross-references: EMBL:AF030052; NID:g2781432; PID:g2781433
A;Experimental source: subsp. Japonica, strain Nipponbare
C;Keywords: glycosyltransferase; hexosyltransferase

```
  Query Match             35.8%;  Score 2118.5;  DB 2;  Length 583;
  Best Local Similarity   66.2%;  Pred. No. 2.4e-154;
  Matches  398;  Conservative   75;  Mismatches   93;  Indels   35;  Gaps    8;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD--PGPKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
              | |:||:|||| |||| |:|| |||  |  || :  |||||||||| ||::  || ||||
Db          1 MAANAGMVAGSRNRNEFVMIRPDGDAPPPAKPGKSVNGQVCQICGDTVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
              ||||||||| ||||||:|| | |||||||||| ||  || |||||: |||||||||:  |
Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRHKGSPRVQGDEEEEDVDDLDNEFNYK-H 119

Qy        119 DSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVD----DIPPEQHALVP 174
               :    |  :       :| | : :  :       :| ||:|| :     |  |::|::
Db        120 GNGKGPEWQI-----QRQGEDVDLSSSSRHEQHRIPRLTSGQQISGEIPDASPDRHSI-- 172

Qy        175 SFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER-MHQ 233
                    |     | |||:||  | :|||||| :||  || |:||: :|: :|:: | |
Db        173 --------RSGTSSYVDPSVPVPVRIVDPSKDLNSYGINSVDWQERVASWRNKQDKNMMQ 224

Qy        234 TGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVL 283
                 |          :| | :|:  |: ::|:||  ||| :|:||:|:| ||::||:||::|
Db        225 VANKYPEARGGDMEGTGSNGE--DIQMVDDARLPLSRIVPIPSNQLNLYRIVIILRLIIL 282
```

```
Qy        284 GFPFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKE 343
              ||| ||| ||| ||: |||:|||||||||:||:|||||||:|| ||||||||:|||:|:|
Db        283 MFFFQYRVTHPVRDAYGLWLVSVICEIWFALSWLLDQFPKWYPINRETYLDRLALRYDRE 342

Qy        344 GQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEAL 403
              |:|||||||| ||||||||||||:| |||||||:||||||||||||||:|||||||
Db        343 GEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVSCYVSDDGSAMLTFEAL 402

Qy        404 SETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVR 463
              |||:|||:||||||||::|||||||||:|| |||||||||:  :||:|||||||||||||
Db        403 SETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKREYEEFKVR 462

Qy        464 INALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVS 523
              |||||||||||||||||| ||| ||||| |||||||||||| ||||| :|||||||||||
Db        463 INALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNELPRLVYVS 522

Qy        524 REKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLG 583
              ||||||: |||||||||||:|||||||  ||||:|||| ||||::|||||||| ||
Db        523 REKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNNSKALREAMCFMMDPALG 582

Qy        584 K 584
              :
Db        583 R 583
```

Search completed: December 15, 2003, 13:04:59
Job time : 33 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:57:19 ; Search time 18 Seconds
(without alignments)
2837.278 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	252	4.3	754	1	BCA1_ACEXY	P19449 acetobacter
2	248	4.2	874	1	BCSA_SALTI	Q8z291 salmonella
3	246	4.2	874	1	BCSA_SALTY	Q93in2 salmonella
4	242	4.1	739	1	BCSA_PSEFL	P58931 pseudomonas
5	239.5	4.0	1596	1	ACS2_ACEXY	Q59167 acetobacter
6	239	4.0	872	1	BCSA_ECO57	Q8x517 escherichia
7	238	4.0	872	1	BCSA_ECOLI	P37653 escherichia
8	237	4.0	756	1	BCA2_ACEXY	O82859 acetobacter
9	230.5	3.9	729	1	BCSA_XANAC	P58932 xanthomonas
10	218	3.7	1518	1	BCA4_ACEXY	Q9rbj2 acetobacter
11	218	3.7	1518	1	BCA5_ACEXY	Q9wx75 acetobacter
12	201	3.4	1550	1	ACS1_ACEXY	P21877 acetobacter
13	194	3.3	745	1	BCA3_ACEXY	Q9wx61 acetobacter
14	139.5	2.4	906	1	CBL_HUMAN	P22681 homo sapien
15	116	2.0	473	1	GLNA_SYNP2	P28605 synechococc
16	113.5	1.9	1254	1	POLS_EEVVE	P36330 venezuelan
17	113	1.9	567	1	COX1_EMENI	P00402 emericella
18	113	1.9	1254	1	POLS_EEVV8	P05674 venezuelan
19	110	1.9	655	1	YS44_CAEEL	Q09372 caenorhabdi
20	110	1.9	1912	1	CHD4_HUMAN	Q14839 homo sapien
21	109	1.8	535	1	COX1_HANWI	P48868 hansenula w
22	109	1.8	896	1	CBL_MOUSE	P22682 mus musculu
23	108.5	1.8	357	1	G6PT_HUMAN	P35575 homo sapien
24	108.5	1.8	1612	1	DNM1_PARLI	Q27746 paracentrot
25	108	1.8	441	1	YCDQ_ECOLI	P75905 escherichia
26	108	1.8	775	1	ICP0_HSV11	P08393 herpes simp
27	107	1.8	1019	1	TOP1_USTMA	P41511 ustilago ma
28	106.5	1.8	992	1	UVRA_MICLU	P13567 micrococcus
29	106.5	1.8	1184	1	FBL2_HUMAN	P98095 homo sapien
30	106	1.8	556	1	SRW1_SCHPO	O13286 schizosacch
31	106	1.8	1254	1	POLS_EEVVT	P09592 venezuelan
32	105.5	1.8	470	1	ROCC_BACSU	P39636 bacillus su
33	105.5	1.8	565	1	YD08_MYCPN	P75472 mycoplasma
34	105	1.8	873	1	COX1_ACACA	Q37370 acanthamoeb
35	105	1.8	1255	1	POLS_EEVV3	P36329 venezuelan
36	104.5	1.8	524	1	COX1_WHEAT	P08741 triticum ae
37	104	1.8	537	1	COX1_SCHPO	P07657 schizosacch
38	103.5	1.7	322	1	NU1M_STRPU	P15548 strongyloce
39	103.5	1.7	524	1	COX1_ORYSA	P14578 oryza sativ
40	103.5	1.7	528	1	COX1_MAIZE	P08742 zea mays (m
41	103.5	1.7	530	1	COX1_SORBI	P05502 sorghum bic
42	103.5	1.7	1120	1	KEFA_ECOLI	P77338 escherichia
43	103.5	1.7	1332	1	XKDO_BACSU	P54334 bacillus su
44	103.5	1.7	4548	1	APOA_HUMAN	P08519 homo sapien
45	103	1.7	677	1	UBF1_XENLA	P25979 xenopus lae

ALIGNMENTS

RESULT 1
BCA1_ACEXY
ID BCA1_ACEXY STANDARD; PRT; 754 AA.
AC P19449;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.
RC STRAIN=1306-3;
RX MEDLINE=91045951; PubMed=2146681;
RA Wong H.C., Fear A.L., Calhoon R.D., Eichinger G.H., Mayer R.,
RA Amikam D., Benziman M., Gelfand D.H., Meade J.H., Emerick A.W.,
RA Bruner R., Ben-Bassat A., Tal R.;
RT "Genetic organization of the cellulose synthase operon in Acetobacter xylinum.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134(1990).
RN [2]
RP ENZYME REGULATION.
RC STRAIN=1306-3;
RX MEDLINE=21194671; PubMed=11297407;
RA Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H.,
RA Volman G., Amikam D., Benziman M., Gilles-Gonzalez M.-A.;
RT "Phosphodiesterase A1, a regulator of cellulose synthesis in Acetobacter xylinum, is a heme-based sensor.";
RL Biochemistry 40:3420-3426(2001).
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC uridine 5'-diphosphate glucose to cellulose in a processive way.
CC The thick cellulosic mats generated by this enzyme probably
CC provide a specialized protective environment to the bacterium.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -!- COFACTOR: Magnesium.
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC (c-di-GMP).
CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to
CC cell growth when O(2) is present.
CC -!- DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC ----------

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M37202; AAA21884.1; -.
DR   PIR; A43735; A43735.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN      147    240       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      317    377       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     26     46       POTENTIAL.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM    108    128       POTENTIAL.
FT   TRANSMEM    167    187       POTENTIAL.
FT   TRANSMEM    407    427       POTENTIAL.
FT   TRANSMEM    430    450       POTENTIAL.
FT   TRANSMEM    468    488       POTENTIAL.
FT   TRANSMEM    516    536       POTENTIAL.
FT   TRANSMEM    549    569       POTENTIAL.
FT   ACT_SITE    189    189       POTENTIAL.
FT   ACT_SITE    333    333       POTENTIAL.
FT   BINDING     236    236       SUBSTRATE (POTENTIAL).
FT   BINDING     238    238       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   754 AA;  84442 MW;  8D5FC1FE62E2C068 CRC64;

  Query Match             4.3%;  Score 252;  DB 1;  Length 754;
  Best Local Similarity   20.4%;  Pred. No. 6.3e-11;
  Matches  168;  Conservative   89;  Mismatches  217;  Indels  350;  Gaps   36;

Qy        275 IIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQ---FPKWFP----- 326
              :|:  : || ||   |         | | ::| :  : : ::| | :   |  |
Db         53 LIVALVCVLVFFIVGRGKSRRTQIF-LEVLSALVSLRY-LTWRLTETLDFDTWIQGGLGV 110

Qy        327 --IERETY-LDRLSLRFDKEGQPSQLAP-----------IDFFVSTVDPLKEPPLVTTN 371
                :  | | |  | | : :  ||   ||          :| |: | |   |   :
Db        111 TLLMAELYALYMLFLSYFQTIQPLHRAPLPLPDNVDDWPTVDIFIPTYD---EQLSIVRL 167

Qy        372 TVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYF 431
              |||  | :|:| |||: |: |||
Db        168 TVLGALGIDWPPDKVNVYILDDG----------------------------------- 190

Qy        432 QQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNN 491
                                  :: |:|:|        ||
Db        191 --------------------VRPEFEQF--------AK---------------------- 200

Qy        492 VRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTN 551
                               ||        :|: |      : | ||| :|  ::     |:
Db        201 ------------------DCGA------LYIGRVD----SSHAKAGNLNHAIK----RTS 228

Qy        552 APYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVF 611
                |:| ||||| |     :: || :|   :  :|:  :| |   |    |
Db        229 GDYILILDCDH-IPTRAFLQIAMGWM---VADRKIALMQTPHHFYSPD------------ 272

Qy        612 FDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNR 671
                                  |:|    ||     : ||                  ||
Db        273 ---------------------PFQRNLAVGY-----RTPPE----------------GN- 289

Qy        672 KQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQS 731
                            ||:                          |::         |  :
Db        290 ---------------LFY--------------------------GVIQD---GNDFWDA 304

Qy        732 SVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGF 791
              : |             | | | :|||  |                |:   :||||  |
Db        305 TFF-------------CGSCAILRREAIESIG-------------GFAVETVTEDAHTAL 338

Qy        792 KMHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG 850
              :|   || : |  ||       | |    |:  : | :||| | |:||  ::  |    |||
Db        339 RMQRRGWSTAYLRIP---VASGLATERLTTHIGQRMRWARGMIQIFRVDNPML----GGG 391

Qy        851 LKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI---TPELNNVASLWFMSLFIC 907
              ||  :|  |::::   : :|| : :   |   |  |: |   :|              :
Db        392 LKLGQRLCYLSAMTSFFFAIPRVIFLASPLAFLFFGQNIIAASP-------------LA 437

Qy        908 IFATSILEMRWS---GVGIDDWWRNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVT 963
              : | :|  |  |      ::  ||   ||  :   :  || |  :: ::      | ||
Db        438 VLAYAIPHMFHSIATAAKVNKGWR-YSFWSEVYETTMALFLVRVTIITLMFPSKGKFNVT 496

Qy        964 SKGG--DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFG--K 1019
                |||  ::|||                          : |   | | |  : | | |  :
Db        497 EKGGVLEEEEFD----------------------LGATYPNIIFAGIMTLGLLIGLFE 532

Qy       1020 LFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWS-----ILLASI 1058
              | | |           | |   |  :  :|:     ||||:|
Db        533 LTFHF----------NQLAGIAKRAYLLNCIWAMISLIILLAAI 566


RESULT 2
BCSA_SALTI
ID   BCSA_SALTI     STANDARD;      PRT;   874 AA.
AC   Q8Z291;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR STY4181 OR T3898.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component for mechanical and chemical protection
CC       at the onset of the stationary phase, when the cells exhibit
CC       multicellular behavior (rdar morphotype). Co-expression of
CC       cellulose and thin aggregative fimbriae leads to a hydrophobic
CC       network with tightly packed cells embedded in a highly inert
CC       matrix (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
```

```
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL627281; CAD08006.1; -.
DR   EMBL; AE016847; AAO71373.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane; Complete proteome.
FT   DOMAIN      271    364       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      441    501       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     30     50       POTENTIAL.
FT   TRANSMEM    151    171       POTENTIAL.
FT   TRANSMEM    173    193       POTENTIAL.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    525    545       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM    592    612       POTENTIAL.
FT   TRANSMEM    634    654       POTENTIAL.
FT   TRANSMEM    668    688       POTENTIAL.
FT   TRANSMEM    833    853       POTENTIAL.
FT   ACT_SITE    313    313       POTENTIAL.
FT   ACT_SITE    457    457       POTENTIAL.
FT   BINDING     360    360       SUBSTRATE (POTENTIAL).
FT   BINDING     362    362       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   874 AA;  100020 MW;  3395D09CD0F51990 CRC64;

  Query Match             4.2%;  Score 248;  DB 1;  Length 874;
  Best Local Similarity   18.0%;  Pred. No. 1.6e-10;
  Matches  152;  Conservative  100;  Mismatches  198;  Indels  394;  Gaps   32;

Qy        258 LSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVN--DAFAL------------W 302
              | |::|   | :    |:|:: | |   :  :|    :|  |  :|            |
Db        189 LVRRMPGRFSAL----MLIVLSLTVSCRYIWWRYTSTLNWDDPVSLVCGLILLFAETYAW 244

Qy        303 LISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPL 362
              :: |:   :| : | |:: |   |: :|                ||   :| || | :
Db        245 IVLVLG--YFQVVWPLNRQP--VPLPKEM--------------SQWPTVDIFVPTYN-- 283

Qy        363 KEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNI 422
              |   |  ||: : | :|:| ||:: :: |||
Db        284 -EDLNVVKNTIYASLGIDWPKDKLNIWILDDGG------------------------- 315

Qy        423 EPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQ 482
                                        :  : :|
Db        316 ------------------------------RESFRQF--------------------- 322

Qy        483 DGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNAL 542
                  :|  |                      |::|      : | ||| :|
Db        323 ------ARHVGVH------------------------YIAR----ATHEHAKAGNINNA 347

Qy        543 VRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID--- 599
              :: :       ::   ||||    |         |: :    |::  :| |  |   |
Db        348 LKHA----KGEFVAIFDCDHVPTRSFLQMTMGWFLKE----KQLAMMQTPHHFFSPDPFE 399

Qy        600 ----RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTC 655
                  |  :  |   :|: :   | |
Db        400 RNLGRFRKTPNEGTLFYGLVQDGND---------------------------------- 424

Qy        656 NCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEK 715
                |    ||  |
Db        425 -MWDATFFCGSC------------------------------------------------ 435

Qy        716 AGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKE 775
                :: :: |:                                                |
Db        436 -AVIRRKPLD-------------------------------------------------E 445

Qy        776 IGWI-YGSVTEDILTGFKMHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSI 833
              || |   :|||| | ::| |: | |  ||:     | |  :||  : | :||| | :
Db        446 IGGIAVETVTEDAHTSLRLHRRGYTSAYMRIPQSA---GLATESLSAHIGQRIRWARGMV 502

Qy        834 EIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPEL 893
              :||  :: ||   :| ||| :|  |:|:: :  : || | : | |   ||   :|
Db        503 QIFRLDN-PL---FGKGLKLAQRLCYLNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---- 554

Qy        894 NNVASLWFMSLFICIFA------TSILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHL 942
                  ::  :| | :|        |:   :  |     :| :|  :     | |   :
Db        555 -----IYAPALMIALFVIPHMVHASLTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT--- 605

Qy        943 FAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLI--PPTTLLLLNFIGVVA 1000
                  |: :|      | ||:|||  ||       |:   :|  |   |:||| :|| |
Db        606 ------LVALINPHKGKFNVTAKGGLVEE-------KYVDWVISRPYIFLVLLNLLGVAA 652

Qy       1001 GVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFS 1060
              ||          |  :|                  :| | |:::          |
Db        653 GV----------WRYYYGP--------------------ENETLTVIV----------S 671

Qy       1061 LLWV 1064
              |:||
Db        672 LVWV 675


RESULT 3
BCSA_SALTY
ID   BCSA_SALTY     STANDARD;      PRT;   874 AA.
AC   Q93IN2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR STM3619.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RX   MEDLINE=21160181; PubMed=11260463;
RA   Zogaj X., Nimtz M., Rohde M., Bokranz W., Roemling U.;
RT   "The multicellular morphotypes of Salmonella typhimurium and
RT   Escherichia coli produce cellulose as the second component of the
RT   extracellular matrix.";
RL   Mol. Microbiol. 39:1452-1463(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=21927388; PubMed=11929533;
RA   Solano C., Garcia B., Valle J., Berasain C., Ghigo J.-M., Gamazo C.,
RA   Lasa I.;
RT   "Genetic analysis of Salmonella enteritidis biofilm formation:
RT   critical role for cellulose.";
RL   Mol. Microbiol. 43:793-808(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
```

```
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component for mechanical and chemical protection
CC       at the onset of the stationary phase, when the cells exhibit
CC       multicellular behavior (rdar morphotype). Co-expression of
CC       cellulose and thin aggregative fimbriae leads to a hydrophobic
CC       network with tightly packed cells embedded in a highly inert
CC       matrix.
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ315770; CAC44015.1; -.
DR   EMBL; AE008867; AAL22479.1; -.
DR   EMBL; AJ315148; CAC86199.1; -.
DR   StyGene; SG?????; bcsA.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane; Complete proteome.
FT   DOMAIN      271    364       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      441    501       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     30     50       POTENTIAL.
FT   TRANSMEM    151    171       POTENTIAL.
FT   TRANSMEM    173    193       POTENTIAL.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    525    545       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM    592    612       POTENTIAL.
FT   TRANSMEM    634    654       POTENTIAL.
FT   TRANSMEM    668    688       POTENTIAL.
FT   TRANSMEM    833    853       POTENTIAL.
FT   ACT_SITE    313    313       POTENTIAL.
FT   ACT_SITE    457    457       POTENTIAL.
FT   BINDING     360    360       SUBSTRATE (POTENTIAL).
FT   BINDING     362    362       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   874 AA;  100043 MW;  4C9421B58606310A CRC64;

  Query Match             4.2%;  Score 246;  DB 1;  Length 874;
  Best Local Similarity  18.1%;  Pred. No. 2.2e-10;

  Matches  153;  Conservative   99;  Mismatches  198;  Indels  394;  Gaps   32;

Qy       258 LSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVN--DAFAL----------W 302
                 | |::|   | :   |:|:: | |   :  :|    :|  |  :|          |
Db       189 LVRRMPGRFSAL----MLIVLSLTVSCRYIWWRYTSTLNWDDPVSLVCGLILLFAETYAW 244

Qy       303 LISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPL 362
             :: |:   :| : | |:: |   |: :|                ||   :| || | :
Db       245 IVLVLG--YFQVVWPLNRQP--VPLPKEM--------------SQWPTVDIFVPTYN-- 283

Qy       363 KEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNI 422
               |   ||: : | :|:| ||:: :: |||
Db       284 -EDLNVVKNTIYASLGIDWPKDKLNIWILDDGGR------------------------- 316

Qy       423 EPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQ 482
                                                 | |:
Db       317 ---------------------------------ESFR--------------------- 320

Qy       483 DGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNAL 542
                   :   | |                      : |::|      : | ||| :|
Db       321 --------HFARHVG--------------------VHYIARTT----HEHAKAGNINNA 347

Qy       543 VRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID--- 599
             :: :        ::   ||||    |        |: :    |::  :| |  |   |
Db       348 LKHA----KGEFVAIFDCDHVPTRSFLQMTMGWFLKE----KQLAMMQTPHHFFSPDPFE 399

Qy       600 ----RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTC 655
                 |  :  |   :|: :   | |
Db       400 RNLGRFRKTPNEGTLFYGLVQDGND---------------------------------- 424

Qy       656 NCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEK 715
               |    ||  |
Db       425 -MWDATFFCGSC------------------------------------------------ 435

Qy       716 AGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKE 775
               :: :: |:                                                |
Db       436 -AVIRRKPLD----------------------------------------------------E 445

Qy       776 IGWI-YGSVTEDILTGFKMHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSI 833
             || |   :|||| | ::| |: | | ||:     | |  :||  : | :||| | :
Db       446 IGGIAVETVTEDAHTSLRLHRRGYTSAYMRIPQAA---GLATESLSAHIGQRIRWARGMV 502

Qy       834 EIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPEL 893
             :||  :: ||   :| |||  :|  |:|:: :  : || | : | |   ||   :|
Db       503 QIFRLDN-PL---FGKGLKLAQRLCYLNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---- 554

Qy       894 NNVASLWFMSLFICIFA------TSILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHL 942
                 ::  :| | :|      |:   :  |     :| :| :     | |   :
Db       555 -----IYAPALMIALFVIPHMVHASLTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT--- 605

Qy       943 FAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLI--PPTTLLLLNFIGVVA 1000
                   |: :|      | ||:||| ||       |:  :|  |   |:||| :|| |
Db       606 ------LVALINPHKGKFNVTAKGGLVEE-------KYVDWVISRPYIFLVLLNLLGVAA 652

Qy      1001 GVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFS 1060
             ||          |   :|                     :| | |:::          |
Db       653 GV----------WRYYYGP--------------------ENETLTVIV----------S 671

Qy      1061 LLWV 1064
             |:||
Db       672 LVWV 675


RESULT 4
BCSA_PSEFL
ID   BCSA_PSEFL     STANDARD;      PRT;   739 AA.
AC   P58931; Q8RSZ1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR WSSB.
OS   Pseudomonas fluorescens.
```

```
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SBW25;
RX   MEDLINE=22013850; PubMed=12019221;
RA   Spiers A.J., Kahn S.G., Bohannon J., Travisano M., Rainey P.B.;
RT   "Adaptive divergence in experimental populations of Pseudomonas
RT   fluorescens. I. Genetic and phenotypic bases of wrinkly spreader
RT   fitness.";
RL   Genetics 161:33-46(2002).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component responsible for the structural
CC       integrity and rigidity of self-supporting mats characteristic of
CC       the "wrinkly spreader" phenotype.
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY074776; AAL71842.1; -.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN      157    250       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      327    387       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     36     55       POTENTIAL.
FT   TRANSMEM     59     76       POTENTIAL.
FT   TRANSMEM     83    101       POTENTIAL.
FT   TRANSMEM    116    138       POTENTIAL.
FT   TRANSMEM    417    436       POTENTIAL.
FT   TRANSMEM    440    462       POTENTIAL.
FT   TRANSMEM    524    546       POTENTIAL.
FT   TRANSMEM    551    573       POTENTIAL.
FT   ACT_SITE    199    199       POTENTIAL.
FT   ACT_SITE    343    343       POTENTIAL.
FT   BINDING     246    246       SUBSTRATE (POTENTIAL).
FT   BINDING     248    248       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   739 AA;  82165 MW;  2B962EA3854B23BB CRC64;

  Query Match             4.1%;  Score 242;  DB 1;  Length 739;
  Best Local Similarity   19.1%;  Pred. No. 3.5e-10;
  Matches  155;  Conservative   86;  Mismatches  207;  Indels  364;  Gaps   33;

Qy        260 RKIPLPSSQINPYRMIIIIRLVV----------------LGF------FFHYRVMHPVN 296
              ||||           : |: |||                |||      || | ::
Db         77 RKIP---------GRLAILALVVLSLVASLRYMFWRLTSTLGFETWVDMFFGYGLV---- 123

Qy        297 DAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFV 356
```

```
Db        124 -AAEFYALIVLIFGYVQTAWPLRRTPVWLKTE----------------PEEWPTVDVFI 165

Qy        357 STVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPF 416
              | :   |   :   |: :  ::|:| ||:  :| |||             :| :    |
Db        166 PTYN---EALSIVKLTIFAAQAMDWPKDKLRVHVLDDG----------RRDDFRE----F 208

Qy        417 CKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPE 476
              |:                    ||   |::|
Db        209 CR---------------------KVGVNYIR---------------------------- 218

Qy        477 EGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKA 536
                                  ||:                                | | ||
Db        219 ----------------RDN-----------------------------------NFHAKA 227

Qy        537 GAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRF- 595
              | :|  ::|    |:  |:   | ||    |        |: ||    |:  :| |  |
Db        228 GNLNEALKV----TDGEYIALFDADHVPTRSFLQVSLGWFLKDP----KLAMLQTPHFFF 279

Qy        596 ------DGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649
                     :|      |   :|: :   | |
Db        280 SPDPFEKNLDTFRAVPNEGELFYGLVQDGND---------------------------- 310

Qy        650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAP 709
                      |   ||
Db        311 -------LWNATFFC---------------------------------------------- 318

Qy        710 GAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDK 769
                                 |  :|     ||| ||          |:
Db        319 -------------------GSCAVIRREPLLEIGGV----------AVE--------- 338

Qy        770 TDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIY-CIPKRVAFKGSAPLNLSDRLHQVLRW 828
                      :||||  |  |::  |: : |  ||:     | |  :||  ::| :||
Db        339 -----------TVTEDAHTALKLNRLGYNTAYLAIPQAA---GLATESLSRHINQRIRW 383

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
              | |  :|| ::: ||    | |||: :|  | |:: : :  :| | : | |   |: |
Db        384 ARGMAQIFRTDN-PL---LGKGLKWGQRICYANAMQHFFYGLPRLVFLTAPLAYLIFG-- 437

Qy        889 ITPELNNVASLWFMSLFI-CIFATSILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHL 942
                |: : ::|  ::  :  :  :|:   |  |     :| || :     | |       |
Db        438 --AEIFHASALMIVAYVLPHLVHSSLTNSRIQGRFRHSFW-NEVYETVLAWYI------L 488

Qy        943 FAVFQGLLKVIAGVDTSFTVTSKGG-DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAG 1001
                |   |:   ||    | || |||  |::|     | |  |  |   || :| ||:  |
Db        489 PPVLVALVNPKAG---GFNVTDKGGIIDKQF-----FDW-KLARPYLVLLAVNLIGLGFG 539

Qy       1002 VSNAINNGYESWGPLFGKLFFAF---WVIVHL 1030
              :  |      ||     :  |    | : :|
Db        540 IHQLI------WGDASTAVTVAINLTWTLYNL 565


RESULT 5
ACS2_ACEXY
ID   ACS2_ACEXY     STANDARD;      PRT;  1596 AA.
AC   Q59167;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase 2 [Includes: Cellulose synthase catalytic subunit
DE   [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose
DE   synthase 2 regulatory domain)].
GN   ACSAII.
OS   Acetobacter xylinus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23769;
RX   MEDLINE=95394846; PubMed=7665515;
RA   Saxena I.M., Brown R.M. Jr.;
RT   "Identification of a second cellulose synthase gene (acsAII) in
```

```
RT   Acetobacter xylinum.";
RL   J. Bacteriol. 177:5276-5283(1995).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- MISCELLANEOUS: It is not essential for cellulose production in
CC       this strain.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC       GLYCOSYLTRANSFERASE FAMILY 2.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15957; AAA85264.1; -.
DR   PIR; T31338; T31338.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR003920; Cellsynth_B.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF03170; BcsB; 1.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
DR   PRINTS; PR01440; CELLSNTHASEB.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN        1    749       CATALYTIC.
FT   DOMAIN      750   1596       CYCLIC DI-GMP BINDING (POTENTIAL).
FT   DOMAIN      145    238       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      315    375       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     25     45       POTENTIAL.
FT   TRANSMEM    106    126       POTENTIAL.
FT   TRANSMEM    396    416       POTENTIAL.
FT   TRANSMEM    421    441       POTENTIAL.
FT   TRANSMEM    505    525       POTENTIAL.
FT   TRANSMEM    544    564       POTENTIAL.
FT   TRANSMEM   1553   1573       POTENTIAL.
FT   ACT_SITE    187    187       POTENTIAL.
FT   ACT_SITE    331    331       POTENTIAL.
FT   BINDING     234    234       SUBSTRATE (POTENTIAL).
FT   BINDING     236    236       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   1596 AA;  175799 MW;  D98A6F6259E1F3CE CRC64;

  Query Match             4.0%;  Score 239.5;  DB 1;  Length 1596;
  Best Local Similarity   18.2%;  Pred. No. 1.6e-09;
  Matches  145;  Conservative   91;  Mismatches  228;  Indels  331;  Gaps   26;

Qy        274 MIIIIRLVVLGFFFHYRVMHPVNDAFALWLIS-------VICEIWFAMSWILDQFPKWFP 326
              ::: :| :|        :  |:  | :| |::        :|  :| ||| ||: |   |
Db         81 LLVSLRYMVWRLTTTLELHSPLQAALSLLLVAAELYALLTLCLSYFQMSWPLDRKP--LP 138

Qy        327 IERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKV 386
              :  :|                :    :| :| : :   |   :  :|||  |::|:| ||:
Db        139 LPADT--------------TDWPVVDVYVPSYN---EELSLVRSTVLGALAIDWPADKL 180

Qy        387 SCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANF 446
              : |: |||
Db        181 NVYILDDG---------------------------------------------------- 188
```

```
Qy        447 VRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQS 506
                  :: :  | :   |                      |  :||
Db        189 ------RRKSFHAFAMEAGA----------------------GYIIRDQ---------- 209

Qy        507 GGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINN 566
                                 |:| ||| :|  :||    |   |::  |||| |
Db        210 ------------------------NNHAKAGNLNHALRV----TEGEYVVIFDCDH-IPT 240

Qy        567 SKAIKEAMCFMM-DPLLGKKVCYVQFPQRFDGIDRHDR--YANRNV-----VFFDINMKG 618
                 :|: : :|| ||    |:  :| |  |   |   |     :||     :|: :   |
Db        241 RGFLKKTIGWMMADP----KLALLQTPHHFYSPDPFQRNLATGQNVPPEGNMFYGLVQDG 296

Qy        619 LDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTK 678
              |      : |:    || |: |
Db        297 NDFWDATFFCGSCAAIRRSAVLGIGG-------------------------------- 322

Qy        679 PKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTST 738
                                                                    | | |
Db        323 ------------------------------------------------------FATET 327

Qy        739 LLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGW 798
                                                          ||||  |  ||   ||
Db        328 ------------------------------------------VTEDAHTALKMQREGW 343

Qy        799 RSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFS 858
                : |   ::    | :   |   : | :||| | ::|   :: ||    | ||:: :|
Db        344 HTAYL--RQPLAAGLSTERLMLHIGQRVRWARGMLQIMRLDN-PL---LGSGLRWQQRLC 397

Qy        859 YINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSL-------FICIFAT 911
              |:::: :   :|| | :   |  |  |:      | :|:  |  |       |  |
Db        398 YLSAMSHFLFAIPRLVFLASPLAFLFLGQ------NIIAASPFAILVYAFPHVFHSIGTL 451

Qy        912 SILEMRWSGVGIDDWWRNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGG-DD 969
              | :| ||              ||  |   :  || |   :: ::       | || |||
Db        452 SRVEGRW----------RYSFWSEIYETTLALFLVRVTIMTLLNPRKGEFNVTDKGGLLQ 501

Qy        970 EEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVH 1029
                |: :|        : |   | ::  : :| |:   :   |     |         || |
Db        502 SEYFDL------NAVYPNVILAVILALALVRGIGGMMWE-YHDRLALQSFALNTLWVAVS 554

Qy       1030 LYPFLKGL-VGRQNR 1043
                |   |  : |||: |
Db        555 LIIVLASIAVGRETR 569
```

```
RESULT 6
BCSA_ECO57
ID   BCSA_ECO57     STANDARD;      PRT;   872 AA.
AC   Q8X5L7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR Z4948 OR ECS4413.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
```

```
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component for mechanical and chemical protection
CC       at the onset of the stationary phase, when the cells exhibit
CC       multicellular behavior (rdar morphotype). Co-expression of
CC       cellulose and thin aggregative fimbriae leads to a hydrophobic
CC       network with tightly packed cells embedded in a highly inert
CC       matrix (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE005579; AAG58675.1; ALT_INIT.
DR   EMBL; AP002565; BAB37836.1; ALT_INIT.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane; Complete proteome.
FT   DOMAIN      271    364       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      441    501       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     30     50       POTENTIAL.
FT   TRANSMEM    151    171       POTENTIAL.
FT   TRANSMEM    173    193       POTENTIAL.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    525    545       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM    592    612       POTENTIAL.
FT   TRANSMEM    640    660       POTENTIAL.
FT   TRANSMEM    668    688       POTENTIAL.
FT   TRANSMEM    833    853       POTENTIAL.
FT   ACT_SITE    313    313       POTENTIAL.
FT   ACT_SITE    457    457       POTENTIAL.
FT   BINDING     360    360       SUBSTRATE (POTENTIAL).
FT   BINDING     362    362       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   872 AA;  99710 MW;  F3F1A24A2B713FBA CRC64;

  Query Match             4.0%;  Score 239;  DB 1;  Length 872;
  Best Local Similarity  17.9%;  Pred. No. 7.5e-10;
  Matches  147;  Conservative   98;  Mismatches  198;  Indels  380;  Gaps   33;

Qy       274 MIIIIRLVVLGFFFHYRVMHPVNDAFALW--LISVICEI---------WFAMSWILDQFP 322
                 |:|:: | |   :  :|    :|     |   :|::| :          |  :  :|  |
Db       201 MLIVLSLTVSCRYIWWRYTSTLN-----WDDPVSLVCGLILLFAETYAWIVL--VLGYFQ 253

Qy       323 KWFPIERETYLDRLSLRFDKEGQPSQLAPIDPFVSTVDPLKEPPLVTTNTVLSILSVDYP 382
               :|: |:     : |  |     ||    :| || | :   |    | ||: : | :|:|
Db       254 VVWPLNRQP----VPLPKDMSLWPS----VDIFVPTYN---EDLNVVKNTIYASLGIDWP 302

Qy       383 VDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKV 442
             ||:: :: |||
Db       303 KDKLNIWILDDGG---------------------------------------------- 315

Qy       443 AANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVF 502
                       : |: :|          || |
Db       316 ----------REEFRQF----------AQNV---------------------------- 326

Qy       503 LGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDH 562
                         : |::|     : | ||| :|  :: :       ::   ||||
Db       327 --------------GVKYIARTT----HEHAKAGNINNALKYA----KGEFVSIFDCDH 363

Qy       563 YINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID-------RHDRYANRNVVFFDIN 615
                 |         |: :    |::  :| |  |   |       |  :  |   :|: :
Db       364 VPTRSFLQMTVGWFLKE----KQLAMMQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLV 419

Qy       616 MKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKK 675
               | |                                     |    ||  |
Db       420 QDGND-----------------------------------MWDATFFCGSC-------- 435

Qy       676 TTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFV 735
                                                       :: :: |:
Db       436 ----------------------------------------AVIRRKPLD---------- 444

Qy       736 TSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMH 794
                                                      |||  |   :||||  |  ::|
Db       445 ---------------------------------------EIGGIAVETVTEDAHTSLRLH 465

Qy       795 CHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKF 853
               |: | |  ||:     | |  :||  : | :||| | ::||  :: ||     | ||||
Db       466 RRGYTSAYMRIPQAA---GLATESLSAHIGQRIRWARGMVQIFRLDN-PL---TGKGLKF 518

Qy       854 LERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFA--- 910
               :|  |:|:: :  : || | : | |   ||   :|         ::  :| | :|
Db       519 AQRLCYVNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---------IYAPALMIALFVLPH 569

Qy       911 ---TSILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHLFAVFQGLLKVIAGVDTSFTV 962
                   |:   :  |     :| :| :     | |   :          |: :|      | |
Db       570 MIHASLTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT---------LVALINPHKGKFNV 619

Qy       963 TSKGG-DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLF 1021
               |:|||  :||: :     |  :  |   |:||| :||  |:          |   :|
Db       620 TAKGGLVEEEYVD-----W-VISRPYIFLVLLNLVGVAVGI----------WRYFYG--- 660

Qy      1022 FAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
                                        ||       :|  : |::||
Db       661 ----------------------PPT------EMLTVVVSMVWV 675


RESULT 7
BCSA_ECOLI
ID   BCSA_ECOLI     STANDARD;      PRT;   872 AA.
AC   P37653; P37654; P76712; P76713; Q8RSS7;
DT   01-OCT-1994 (Rel. 30, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR B3533.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94316500; PubMed=8041620;
```

```
RA   Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT   "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT   region from 76.0 to 81.5 minutes.";
RL   Nucleic Acids Res. 22:2576-2586(1994).
RN   [2]
RP   REVISIONS.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   CHARACTERIZATION.
RC   STRAIN=ECOR10, ECOR12, and TOB1;
RX   MEDLINE=21160181; PubMed=11260463;
RA   Zogaj X., Nimtz M., Rohde M., Bokranz W., Roemling U.;
RT   "The multicellular morphotypes of Salmonella typhimurium and
RT   Escherichia coli produce cellulose as the second component of the
RT   extracellular matrix.";
RL   Mol. Microbiol. 39:1452-1463(2001).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component for mechanical and chemical protection
CC       at the onset of the stationary phase, when the cells exhibit
CC       multicellular behavior (rdar morphotype). Co-expression of
CC       cellulose and thin aggregative fimbriae leads to a hydrophobic
CC       network with tightly packed cells embedded in a highly inert
CC       matrix.
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains the
CC       conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- MISCELLANEOUS: The genes bscA, bcsB, bcsZ and bcsC are
CC       constitutively transcribed but cellulose synthesis occurs only
CC       when adrA, a putative transmembrane protein regulated by agfD, is
CC       expressed. Cellulose production is abolished in E.coli K12.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 128.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00039; AAB18510.1; ALT_FRAME.
DR   EMBL; U00039; AAB18511.1; ALT_FRAME.
DR   EMBL; AE000430; AAC76558.1; ALT_INIT.
DR   EcoGene; EG12260; bcsA.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane; Complete proteome.
```

```
FT   DOMAIN      271    364       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      441    501       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     30     50       POTENTIAL.
FT   TRANSMEM    151    171       POTENTIAL.
FT   TRANSMEM    173    193       POTENTIAL.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    525    545       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM    592    612       POTENTIAL.
FT   TRANSMEM    640    660       POTENTIAL.
FT   TRANSMEM    668    688       POTENTIAL.
FT   TRANSMEM    833    853       POTENTIAL.
FT   ACT_SITE    313    313       POTENTIAL.
FT   ACT_SITE    457    457       POTENTIAL.
FT   BINDING     360    360       SUBSTRATE (POTENTIAL).
FT   BINDING     362    362       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   872 AA;  99784 MW;  14326B8A2EB228F7 CRC64;

  Query Match             4.0%;  Score 238;  DB 1;  Length 872;
  Best Local Similarity  17.9%;  Pred. No. 8.9e-10;
  Matches  147;  Conservative   98;  Mismatches  198;  Indels  380;  Gaps   33;

Qy        274 MIIIIRLVVLGFFFHYRVMHPVNDAFALW--LISVICEI---------WFAMSWILDQFP 322
              |:|:: | |   :  :|    :|     |   :|::| :         |  :  :|  |
Db        201 MLIVLSLTVSCRYIWWRYTSTLN-----WDDPVSLVCGLILLFAETYAWIVL--VLGYFQ 253

Qy        323 KWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYP 382
                :|: |:     : |  |    ||    :| || | :   |   |  ||: : | :|:|
Db        254 VVWPLNRQP----VPLPKDMSLWPS----VDIFVPTYN---EDLNVVKNTIYASLGIDWP 302

Qy        383 VDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKV 442
              ||:: :: |||
Db        303 KDKLNIWILDDGG---------------------------------------------- 315

Qy        443 AANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVF 502
                    : |: :|          || |
Db        316 ----------REEFRQF----------AQNV---------------------------- 326

Qy        503 LGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDH 562
                         : |::|      : | ||| :|  :: :       ::    ||||
Db        327 --------------GVKYIARTT----HEHAKAGNINNALKYA----KGEFVSIFDCDH 363

Qy        563 YINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID-------RHDRYANRNVVFFDIN 615
                  |         |: :    |::  :| |  |   |       |  :  |   :|: :
Db        364 VPTRSFLQMTMGWFLKE----KQLAMMQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLV 419

Qy        616 MKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKK 675
              | |                                   |   ||  |
Db        420 QDGND-----------------------------------MWDATFFCGSC-------- 435

Qy        676 TTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFV 735
                                                      :: :: |:
Db        436 ---------------------------------------AVIRRKPLD---------- 444

Qy        736 TSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMH 794
                                                    |||  |   :||||  |  ::|
Db        445 -------------------------------------EIGGIAVETVTEDAHTSLRLH 465

Qy        795 CHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKF 853
              |: | |  ||:     | |  :||  : | :||| | ::||  :: ||    | ||||
Db        466 RRGYTSAYMRIPQAA---GLATESLSAHIGQRIRWARGMVQIFRLDN-PL---TGKGLKF 518

Qy        854 LERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFA--- 910
             :|  |:|:: :  : || | : | |   ||   :|         ::  :| | :|
Db        519 AQRLCYVNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---------IYAPALMIALFVLPH 569

Qy        911 ---TSILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHLFAVFQGLLKVIAGVDTSFTV 962
                 |:   :  |     :| :|  :     | |   :          |: :|       | |
Db        570 MIHASLTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT---------LVALINPHKGKFNV 619

Qy        963 TSKGG-DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLF 1021
              |:|||  :||: :      |  :  |   |:||| :||  |:        |   :|
Db        620 TAKGGLVEEEYVD-----W-VISRPYIFLVLLNLVGVAVGI----------WRYFYG--- 660
```

```
Qy       1022 FAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
                                  ||       :|  : |::||
Db        661 ----------------------PPT------EMLTVVVSMVWV 675


RESULT 8
BCA2_ACEXY
ID   BCA2_ACEXY     STANDARD;      PRT;   756 AA.
AC   O82859;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA.
OS   Acetobacter xylinus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BPR 2001;
RX   MEDLINE=98296257; PubMed=9630539;
RA   Nakai T., Moriya A., Tonouchi N., Tsuchida T., Yoshinaga F.,
RA   Horinouchi S., Sone Y., Mori H., Sakai F., Hayashi T.;
RT   "Control of expression by the cellulose synthase (bcsA) promoter
RT   region from Acetobacter xylinum BPR 2001.";
RL   Gene 213:93-100(1998).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose. The thick cellulosic
CC       mats generated by this enzyme probably provide a specialized
CC       protective environment to the bacterium (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB010645; BAA31463.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN      147    242       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      319    379       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     27     47       POTENTIAL.
FT   TRANSMEM     49     69       POTENTIAL.
FT   TRANSMEM    106    126       POTENTIAL.
FT   TRANSMEM    167    187       POTENTIAL.
FT   TRANSMEM    409    429       POTENTIAL.
FT   TRANSMEM    432    452       POTENTIAL.
FT   TRANSMEM    470    490       POTENTIAL.
FT   TRANSMEM    517    537       POTENTIAL.
FT   TRANSMEM    551    571       POTENTIAL.
FT   ACT_SITE    189    189       POTENTIAL.
FT   ACT_SITE    335    335       POTENTIAL.
FT   BINDING     238    238       SUBSTRATE (POTENTIAL).
FT   BINDING     240    240       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   756 AA;  84561 MW;  6954F39A25E73B0A CRC64;

  Query Match             4.0%;  Score 237;  DB 1;  Length 756;
  Best Local Similarity   19.2%;  Pred. No. 8.6e-10;
  Matches  155;  Conservative   96;  Mismatches  247;  Indels  308;  Gaps   31;

Qy        269 INPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQ---FPKWF 325
              :|    :|:  : |: ||   |         | | ::| :  : : ::| | :   |  |
Db         47 LNNNEQLIVAAVCVVIFFVVGRGKSRRTQIF-LEVLSALVSLRY-LTWRLTETLDFNTWI 104

Qy        326 P-------IERETY-LDRLSLRFDKEGQPSQLAP-----------IDFFVSTVDPLKEP 365
                      :  | | |  | | : :  ||   ||           :| |: | |   |
Db        105 QGILGVILLMAELYALYMLFLSYFQTIQPLHRAPLPLPDNVDDWPTVDIFIPTYD---EQ 161

Qy        366 PLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPR 425
              :  |||  | :|:| |||: |: |||
Db        162 LSIVRLTVLGALGIDWPPDKVNVYILDDG------------------------------ 190

Qy        426 APEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGT 485
                                    :: |:|:|     ||
Db        191 --------------------------VRPEFEQFAKDCGALYI--------------- 207

Qy        486 PWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRV 545
                                  | :| :                : | ||| :|  ::
Db        208 --------------------GRVDVD-----------------SAHAKAGNLNHAIK- 227

Qy        546 SAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYA 605
                |:  |:| ||||| |     :: || :|   :  :|:  :| |  |   |
Db        228 ---RTSGDYILILDCDH-IPTRAFLQIAMGWM---VADRKIALMQTPHHFYSPD------ 274

Qy        606 NRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCC 665
                                          |:|    ||     : ||
Db        275 --------------------------PFQRNLAVGY-----RTPPE------------- 289

Qy        666 CCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLE 725
                ||                 ||:                              |::
Db        290 ---GN----------------LFY--------------------------GVIQD---G 300

Qy        726 KKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTE 785
                |  :: |                | | | :|||  |                |:   :|||
Db        301 NDFWDATFF-------------CGSCAILRREAIESIG-------------GFAVETVTE 334

Qy        786 DILTGFKMHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLW 844
              |  |  :|   || : |  ||       | |   |:  : | :|||  | |:||  ::  |
Db        335 DAHTALRMQRRGWSTAYLRIP---VASGLATERLTTHIGQRMRWARGMIQIFRVDNPML- 390

Qy        845 YGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSL 904
                | |||  :|  |::::   : :|| : :   |   |  |: |        : : :
Db        391 ---GRGLKLGQRLCYLSAMTSFFFAIPRVIFLASPLAFLFAGQNIIAAAPLAVAAYALPH 447

Qy        905 FICIFATSILEMRWSGVGIDDWWRNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVT 963
                  ||:          ::  ||   ||  :   :  || |   :: ::      | ||
Db        448 MFHSIATA--------AKVNKGWR-YSFWSEVYETTMALPLVRVTIVTLLFPSKGKFNVT 498

Qy        964 SKGG--DDEEFSELYTFK---WTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFG 1018
              |||  ::|||    |:    : |:::   : |   |    :    | ||
Db        499 EKGGVLEEEEFDLGATYPNIIFATIMMGGLLIGLFELIVRFNQLDVIARNAY-------- 550

Qy       1019 KLFFAFWVIVHLYPFLKGL-VGRQNR 1043
                |    | :: |      : |||: :
Db        551 -LLNCAWALISLIILFAAIAVGRETK 575


RESULT 9
BCSA_XANAC
ID   BCSA_XANAC     STANDARD;      PRT;   729 AA.
```

```
AC   P58932;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSA OR XAC3518.
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose, which is produced as
CC       an extracellular component for mechanical and chemical protection
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the protein: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE012000; AAM38361.1; -.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane; Complete proteome.
FT   DOMAIN      151     244       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      321     381       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     30      50       POTENTIAL.
FT   TRANSMEM    110     130       POTENTIAL.
FT   TRANSMEM    171     191       POTENTIAL.
FT   TRANSMEM    405     425       POTENTIAL.
FT   TRANSMEM    427     447       POTENTIAL.
FT   TRANSMEM    520     540       POTENTIAL.
FT   TRANSMEM    549     569       POTENTIAL.
FT   TRANSMEM    610     630       POTENTIAL.
FT   ACT_SITE    193     193       POTENTIAL.
FT   ACT_SITE    337     337       POTENTIAL.
FT   BINDING     240     240       SUBSTRATE (POTENTIAL).
FT   BINDING     242     242       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   729 AA;  80915 MW;  B9C08BB995E795B1 CRC64;

  Query Match             3.9%;  Score 230.5;  DB 1;  Length 729;
  Best Local Similarity   18.9%;  Pred. No. 2.5e-09;
  Matches  139;  Conservative   82;  Mismatches  185;  Indels  331;  Gaps   26;

Qy        299 FALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVST 358
              | | |  :  |::  :  :|  |   :|: |:     : |  |:   ||    :| |: |
Db        110 FILGLGLLGAELYAFVILVLGYFQVLWPLNRKP----VPLPADQRLWPS----VDVFIPT 161

Qy        359 VDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCK 418
               :   ||  |   |||:   :|:|  |:: :: |||
Db        162 YN---EPLSVVRTTVLAASVIDWPAGKITIHLLDDG---------------------- 194

Qy        419 RYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEG 478
                                          ||      :||:
Db        195 ---------------------------------RR------DEFR---------------- 200

Qy        479 WTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGA 538
                                        |  : |           : ||:|      | | |||
Db        201 ----------------------AFCAEVG----------INYVTRTN----NAHAKAGN 223

Qy        539 MNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAM-CFMMDPLLGKKVCYVQFPQRF-- 595
               :|| ::      :  |:   |||| |     :: ||  |: |     |:  || |  |
Db        224 INAALK----KCSGDYVAIFDCDH-IPTRSFLQVAMGWFLHD----TKLALVQMPHYFFS 274

Qy        596 -----DGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
                     :| | :  |   :|: :   | |       : |:  | :| ||
Db        275 PDPFERNLDTHGKVPNEGELFYGLLQDGNDQWNATFFCGSCAVIKRTAL---------- 323

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPG 710

Db        324 ---------------------------------------------------------- 323

Qy        711 AENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKT 770

Db        324 ---------------------------------------------------------- 323

Qy        771 DWGKEIGWI-YGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWA 829
                 :|:| :   :||||  |  |:   |:|: |    : |  | |  :||  : | :|||
Db        324 ---EEVGGVAVETVTEDAHTALKLQRRGYRTAYLAVPQAA--GLATESLSGHVAQRIRWA 378

Qy        830 LGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI 889
                | :|   :: ||    | |||  :|  |:|:::: :  :| : | | |   |  |  :
Db        379 RGMAQIARIDN-PL---LGRGLKLSQRLCYLNAMLHFFYGVPRIIYLTAPLAYLFFGAHV 434

Qy        890 TPELNNVASLWFMSLFICIFA-TSILEMRWSGVGIDDWWR----NEQF-----WVIGGVS 939
                      :   :| |  :|   ||:   : : :   :|    || :     | |
Db        435 ---------IQASALMILAYALPHILQANLTNLRVQSRFRHLLWNEVYETTLAWYI---- 481

Qy        940 SHLFAVFQGLLKVIAGVDT-SFTVTSKGG------DDEEFSELYTFKWTTLLIPPTTLLL 992
                    |:  |  :        | || |||        | : :: | |            |||
Db        482 ------FRPTLVALLNPKLGKFNVTPKGGLVARSYFDAQIAKPYLF----------LLL 524

Qy        993 LNFIGVVAGVSNAINNG 1009
              || :|:||||   |  |
Db        525 LNVVGMVAGVLRLIYVG 541


RESULT 10
BCA4_ACEXY
ID   BCA4_ACEXY     STANDARD;      PRT;  1518 AA.
AC   Q9RBJ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative cellulose synthase 2 [Includes: Cellulose synthase
DE   catalytic subunit [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding
DE   domain (Cellulose synthase 2 regulatory subunit)].
GN   BCSABII-A.
OS   Acetobacter xylinus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 7664 / IFO 13693;
RX   MEDLINE=99310341; PubMed=10382968;
RA   Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA   Inoue Y.;
RT   "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT   7664: implication of a novel set of cellulose synthase genes.";
RL   DNA Res. 6:109-115(1999).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC       GLYCOSYLTRANSFERASE FAMILY 2.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB015803; BAA77593.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR003920; Cellsynth_B.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF03170; BcsB; 1.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
DR   PRINTS; PR01440; CELLSNTHASEB.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN        1    731       CATALYTIC.
FT   DOMAIN      732   1518       CYCLIC DI-GMP BINDING (BY SIMILARITY).
FT   DOMAIN      144    237       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      314    374       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     24     44       POTENTIAL.
FT   TRANSMEM     71     91       POTENTIAL.
FT   TRANSMEM    105    125       POTENTIAL.
FT   TRANSMEM    404    424       POTENTIAL.
FT   TRANSMEM    427    447       POTENTIAL.
FT   TRANSMEM    465    485       POTENTIAL.
FT   TRANSMEM    514    534       POTENTIAL.
FT   TRANSMEM    543    563       POTENTIAL.
FT   TRANSMEM   1481   1501       POTENTIAL.
FT   ACT_SITE    186    186       POTENTIAL.
FT   ACT_SITE    330    330       POTENTIAL.
FT   BINDING     233    233       SUBSTRATE (POTENTIAL).
FT   BINDING     235    235       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   1518 AA;  166404 MW;  0EC99B35B6DE4543 CRC64;
```

```
  Query Match             3.7%;  Score 218;  DB 1;  Length 1518;
  Best Local Similarity   18.0%;  Pred. No. 6.4e-08;
  Matches  145;  Conservative   88;  Mismatches  232;  Indels  340;  Gaps   25;

Qy        275 IIIIRLVVLGFFFHYRVMHPVNDAFALWL------ISVICEIWFAMSWILDQFPKWFPIE 328
              :::: |:|   :  :|:   |   |: ||      : :: | :  |:  |  |   :|:
Db         75 LMMLSLLVSLRYIVWRLTATVQ--FSNWLQTALAVLLLLAEAYALMTLCLSYFQMAWPLR 132

Qy        329 RETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSC 388
              |  :    |  |    ||    :| || : :   |   :  :|||  | :|:| |:::
Db        133 RREH----PLPEDMAQWPS----VDVFVPSYN---EELSLVRSTVLGALDLDWPADRLNV 181

Qy        389 YVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVR 448
              |: |||
Db        182 YILDDG---------------------------------------------------- 187

Qy        449 ERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGG 508
                  :: : :| |   |
Db        188 ----RRKAFHDFAVEAGA---------------------------------------- 201

Qy        509 LDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSK 568
                        |: | :    |:| ||| :|  : |    |::|: :  ||||
Db        202 ----------GYIIRAE----NNHAKAGNLNHALAV----TDSPFAVIFDCDHVPTRGF 242

Qy        569 AIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYV 628
                :     | ||     :  :| |  |   |   |            |: |  |:  |
Db        243 LRRTIGWMMADP----NLALLQTPHHFYAPDPFQR----------NLAG--GMHVP--- 282

Qy        629 GTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLF 688
                | :|      | |              |     ||  |
Db        283 PEGNMFYGLVQDGND------------FWDATFFCGSC--------------------- 308

Qy        689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748

Db        309 ---------------------------------------------------------- 308

Qy        749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
                  | : :||:  |              |:   :||||  |  ||   || : |
Db        309 ---AIIRREAVMGIG------------GFATETVTEDAHTALKMQRRGWGTAYLREPLA 352

Qy        809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
              |  | |   |   : | :||| | |:|   ::  |    | ||:: :|  |:::: :
Db        353 A--GLATERLILHIGQRVRWARGMIQIMRLDNPML----GAGLRWEQRLCYLSAMSHFLF 406

Qy        869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFM-SLFICIFATSILEMRWSGVGIDDWW 927
              :|| | :   |   |  |: |          |:: :  :|  :   | :| ||
Db        407 AIPRLTFLVSPLAFLFLGQNIIAASPLAISVYALPHIFHSVITLSRIEGRW--------- 457

Qy        928 RNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIP 986
                  ||  |   |  || |   :: ::      | || |||
Db        458 -RYSFWSEIYETSLALFLVRITIVTLLQPHKGKFNVTDKGG------------------ 497

Qy        987 PTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPT 1046
                              :  ||  |  ::  :  |  :   |   |:|: |
Db        498 ------------------LLARGYFDWDAVYPNVILAGVLCAAL---LRGVFG------- 529

Qy       1047 IVIVWSI-----LLASIFSLLWVRI 1066
                |||       | : | : ||| |
Db        530 --IVWQFHDRLALQSFILNTLWVVI 552


RESULT 11
BCA5_ACEXY
ID   BCA5_ACEXY     STANDARD;      PRT;  1518 AA.
AC   Q9WX75;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative cellulose synthase 3 [Includes: Cellulose synthase
DE   catalytic subunit [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding
DE   domain (Cellulose synthase 3 regulatory subunit)].
GN   BCSABII-B.
OS   Acetobacter xylinus.
```

```
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 7664 / IFO 13693;
RX   MEDLINE=99310341; PubMed=10382968;
RA   Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA   Inoue Y.;
RT   "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT   7664: implication of a novel set of cellulose synthase genes.";
RL   DNA Res. 6:109-115(1999).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC       GLYCOSYLTRANSFERASE FAMILY 2.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB015804; BAA77600.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR003920; Cellsynth_B.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF03170; BcsB; 1.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
DR   PRINTS; PR01440; CELLSNTHASEB.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN        1    731       CATALYTIC.
FT   DOMAIN      732   1518       CYCLIC DI-GMP BINDING (BY SIMILARITY).
FT   DOMAIN      144    237       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      314    374       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     24     44       POTENTIAL.
FT   TRANSMEM     71     91       POTENTIAL.
FT   TRANSMEM    105    125       POTENTIAL.
FT   TRANSMEM    404    424       POTENTIAL.
FT   TRANSMEM    428    448       POTENTIAL.
FT   TRANSMEM    465    485       POTENTIAL.
FT   TRANSMEM    514    534       POTENTIAL.
FT   TRANSMEM    543    563       POTENTIAL.
FT   TRANSMEM   1481   1501       POTENTIAL.
FT   ACT_SITE    186    186       POTENTIAL.
FT   ACT_SITE    330    330       POTENTIAL.
FT   SITE        233    233       IMPORTANT FOR SUBSTRATE BINDING
FT                                (POTENTIAL).
FT   SITE        235    235       IMPORTANT FOR SUBSTRATE BINDING
FT                                (POTENTIAL).
SQ   SEQUENCE   1518 AA;  166464 MW;  7D7634503183DAB6 CRC64;

  Query Match             3.7%;  Score 218;  DB 1;  Length 1518;
  Best Local Similarity   18.0%;  Pred. No. 6.4e-08;
  Matches  145;  Conservative   88;  Mismatches  232;  Indels  340;  Gaps   25;
```

```
Qy      275 IIIIRLVVLGFFFHYRVMHPVNDAFALWL------ISVICEIWFAMSWILDQFPKWFPIE 328
            :::: |:|   :  :|:   |   |: ||      : :: | :  |:  |  |   :|:
Db       75 LMMLSLLVSLRYIVWRLTATVQ--FSNWLQTALAVLLLLAEAYALMTLCLSYFQMAWPLR 132

Qy      329 RETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSC 388
            |  :     |  |    ||    :| || : :   |   :  :|||  | :|:| |:::
Db      133 RREH----PLPEDMAQWPS----VDVFVPSYN---EELSLVRSTVLGALDLDWPADRLNV 181

Qy      389 YVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVR 448
            |: |||
Db      182 YILDDG---------------------------------------------------- 187

Qy      449 ERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGG 508
                :: : :| |   |
Db      188 ----RRKAFHDFAVEAGA------------------------------------------ 201

Qy      509 LDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSK 568
                      |: | :    |:| ||| :|  : |    |::|: :  ||||
Db      202 -----------GYIIRAE----NNHAKAGNLNHALAV----TDSPFAVIFDCDHVPTRGF 242

Qy      569 AIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYV 628
              :     | ||     :  :| | |   |   |           |: |  |:  |
Db      243 LRRTIGWMMADP----NLALLQTPHHFYAPDPFQR----------NLAG--GMHVP--- 282

Qy      629 GTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLF 688
            | :|      | |              |    ||  |
Db      283 PEGNMFYGLVQDGND----------FWDATFFCGSC---------------------- 308

Qy      689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748
Db      309 ---------------------------------------------------------- 308

Qy      749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
              | : :||:  |             |:   :|||| |  ||   || : |
Db      309 ---AIIRREAVMGIG----------GFATETVTEDAHTALKMQRRGWGTAYLREPLA 352

Qy      809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
            |  | |   |   : | :||| | |:|   ::  |    | ||:: :|  |:::: :
Db      353 A--GLATERLILHIGQRVRWARGMIQIMRLDNPML----GAGLRWEQRLCYLSAMSHFLF 406

Qy      869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFM-SLFICIFATSILEMRWSGVGIDDWW 927
            :|| | :   |   | |: |        |:: :  :|  :   | :| ||
Db      407 AIPRLTFLVSPLAFLFLGQNIIAASPLAISVYALPHIFHSVITLSRIEGRW--------- 457

Qy      928 RNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIP 986
              ||  |   |  || |   :: ::       | || |||
Db      458 -RYSFWSEIYETSLALFLVRITIVTLLQPHKGKFNVTDKGG------------------- 497

Qy      987 PTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPT 1046
                                  :  ||  |  ::  :  |  :   |   |:|: |
Db      498 -----------------LLARGYFDWDAVYPNVILAGVLCAAL---LRGVFG------- 529

Qy     1047 IVIVWSI-----LLASIFSLLWVRI 1066
              |||       | : | : ||| |
Db      530 --IVWQFHDRLALQSFILNTLWVVI 552
```

```
RESULT 12
ACS1_ACEXY
ID   ACS1_ACEXY     STANDARD;      PRT;  1550 AA.
AC   P21877; P37717;
DT   01-MAY-1991 (Rel. 18, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cellulose synthase 1 [Includes: Cellulose synthase catalytic domain
DE   [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose
DE   synthase 1 regulatory domain)].
GN   ACSAB OR ACSA OR ACSB.
OS   Acetobacter xylinus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
```

```
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 6-20.
RC   STRAIN=ATCC 53582;
RX   MEDLINE=91346705; PubMed=2151718;
RA   Saxena I.M., Lin F.C., Brown R.M. Jr.;
RT   "Cloning and sequencing of the cellulose synthase catalytic subunit
RT   gene of Acetobacter xylinum.";
RL   Plant Mol. Biol. 15:673-683(1990).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 768-781.
RC   STRAIN=ATCC 53582;
RX   MEDLINE=91322509; PubMed=1830823;
RA   Saxena I.M., Lin F.C., Brown R.M. Jr.;
RT   "Identification of a new gene in an operon for cellulose biosynthesis
RT   in Acetobacter xylinum.";
RL   Plant Mol. Biol. 16:947-954(1991).
RN   [3]
RP   REVISIONS.
RC   STRAIN=ATCC 53582;
RX   MEDLINE=94364954; PubMed=8083166;
RA   Saxena I.M., Kudlicka K., Okuda K., Brown R.M. Jr.;
RT   "Characterization of genes in the cellulose-synthesizing operon (acs
RT   operon) of Acetobacter xylinum: implications for cellulose
RT   crystallization.";
RL   J. Bacteriol. 176:5735-5752(1994).
RN   [4]
RP   SEQUENCE OF 1-8 FROM N.A.
RC   STRAIN=ATCC 23769;
RX   MEDLINE=94131945; PubMed=8300521;
RA   Standal R., Iversen T.-G., Coucheron D.H., Fjaervik E., Blatny J.M.,
RA   Valla S.;
RT   "A new gene required for cellulose production and a gene encoding
RT   cellulolytic activity in Acetobacter xylinum are colocalized with the
RT   bcs operon.";
RL   J. Bacteriol. 176:665-672(1994).
RN   [5]
RP   FUNCTION.
RC   STRAIN=ATCC 53582;
RX   MEDLINE=90202941; PubMed=2138620;
RA   Lin F.C., Brown R.M. Jr., Drake R.R. Jr., Haley B.E.;
RT   "Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding
RT   subunit of cellulose synthase in Acetobacter xylinum using the
RT   photoaffinity probe 5-azido-UDP-Glc.";
RL   J. Biol. Chem. 265:4782-4784(1990).
RN   [6]
RP   3D-STRUCTURE MODELING, AND MUTAGENESIS OF ASP-188; ASP-189; ASP-236;
RP   ASP-333; GLN-369; ARG-370 AND ARG-372.
RC   STRAIN=ATCC 23769;
RX   MEDLINE=21324818; PubMed=11430986;
RA   Saxena I.M., Brown R.M. Jr., Dandekar T.;
RT   "Structure-function characterization of cellulose synthase:
RT   relationship to other glycosyltransferases.";
RL   Phytochemistry 57:1135-1148(2001).
RN   [7]
RP   REVIEW ON DOMAIN ARCHITECTURE.
RX   MEDLINE=95189716; PubMed=7883697;
RA   Saxena I.M., Brown R.M. Jr., Fevre M., Geremia R.A., Henrissat B.;
RT   "Multidomain architecture of beta-glycosyl transferases: implications
RT   for mechanism of action.";
RL   J. Bacteriol. 177:1419-1424(1995).
CC   -!- FUNCTION: Bifunctional protein comprised of a catalytic subunit
CC       and a regulatory subunit. The catalytic subunit of cellulose
CC       synthase polymerizes uridine 5'-diphosphate glucose to cellulose
CC       in a processive way. The thick cellulosic mats generated by this
CC       enzyme probably provide a specialized protective environment to
CC       the bacterium. The regulatory subunit binds bis-(3'-5') cyclic
CC       diguanylic acid (c-di-GMP).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by c-di-GMP.
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- INDUCTION: Cellulose is produced at a linear rate with respect to
CC       cell growth when O(2) is present.
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC       GLYCOSYLTRANSFERASE FAMILY 2.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC       FAMILY.
CC   -!- CAUTION: Was originally (Ref.1) thought to be two separate ORFs
CC       named acsA and acsB, due to a frameshift in position 678.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54676; CAA38487.1; ALT_FRAME.
DR   EMBL; X54676; CAA38488.1; ALT_FRAME.
DR   EMBL; M96060; AAA16971.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF03170; BcsB; 1.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN        1    741       CATALYTIC.
FT   DOMAIN      742   1550       CYCLIC DI-GMP BINDING.
FT   DOMAIN      147    240       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      317    377       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     26     46       POTENTIAL.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM    106    126       POTENTIAL.
FT   TRANSMEM    398    418       POTENTIAL.
FT   TRANSMEM    423    443       POTENTIAL.
FT   TRANSMEM    468    488       POTENTIAL.
FT   TRANSMEM    507    527       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM   1513   1533       POTENTIAL.
FT   ACT_SITE    189    189       POTENTIAL.
FT   ACT_SITE    333    333       POTENTIAL.
FT   BINDING     236    236       SUBSTRATE (POTENTIAL).
FT   BINDING     238    238       SUBSTRATE (POTENTIAL).
FT   MUTAGEN     188    188       D->N: DECREASE IN ACTIVITY.
FT   MUTAGEN     188    188       D->P: LOSS OF ACTIVITY.
FT   MUTAGEN     189    189       D->Y: LOSS OF ACTIVITY.
FT   MUTAGEN     236    236       D->Y: LOSS OF ACTIVITY.
FT   MUTAGEN     333    333       D->R: LOSS OF ACTIVITY.
FT   MUTAGEN     369    369       Q->M: LOSS OF ACTIVITY.
FT   MUTAGEN     370    370       R->P: LOSS OF ACTIVITY.
FT   MUTAGEN     370    370       R->Q: DECREASE IN ACTIVITY.
FT   MUTAGEN     372    372       R->A: LOSS OF ACTIVITY.
SQ   SEQUENCE   1550 AA;  168161 MW;  63AB8952BC39E961 CRC64;

  Query Match             3.4%;  Score 201;  DB 1;  Length 1550;
  Best Local Similarity   19.8%;  Pred. No. 1.3e-06;
  Matches  156;  Conservative   85;  Mismatches  212;  Indels  334;  Gaps   37;

Qy        298 AFALWL------ISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAP 351
              :|  ||      : :: |::  |   |  |    |: |                 | :
Db         99 SFDTWLQGLLGTMLLVAELYALMMLFLSYFQTIAPLHRAP--------LPLPPNPDEWPT 150

Qy        352 IDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAK 411
              :| || | :   |   :   |||  | :|:| :||  :: |||
```

```
Db        151 VDIFVPTYN---EELSIVRLTVLGSLGIDWPPEKVRVHILDDG---------------- 190

Qy        412 KWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKA 471
                          | ||:              | |  |
Db        191 ------------RRPEF-------------AAFAAE---------------------- 201

Qy        472 QKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYN 531
                                                  |  |      |::|      |
Db        202 ----------------------------------------CGAN------YIARPT----N 212

Qy        532 HHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMM-DPLLGKKVCYVQ 590
               | ||| :|  :      |:  |:|  |||| :     ::  | :|: ||    |:  :|
Db        213 EHAKAGNLNYAIG----HTDGDYILIFDCDH-VPTRAFLQLTMGWMVEDP----KIALMQ 263

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
               |  |   |                                  |:|    ||     : |
Db        264 TPHHFYSPD---------------------------------PFQRNLSAGY-----RTP 285

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPG 710
               |                ||                ||:
Db        286 PE---------------GN---------------LFY---------------------- 292

Qy        711 AENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKT 770
                 |:|        |  :: |              | | | : ||  |
Db        293 ------GVVQD---GNDFWDATFF-------------CGSCAILRRTAIEQIG------- 323

Qy        771 DWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGS-APLNLSDRLHQVLRWA 829
                 |:   :||||  |  ||   || : |    |:   |  |   |   : | :|||
Db        324 ------GFATQTVTEDAHTALKMQRLGWSTAYL---RIPLAGGLATERLILHIGQRVRWA 374

Qy        830 LGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI 889
               | ::||  :: ||   :| || : :|  |::::     ::| : : : |   |  |: |
Db        375 RGMLQIFRIDN-PL---FGRGLSWGQRLCYLSAMTSFLFAVPRVIFLSSPLAFLFFGQNI 430

Qy        890 ---TPELNNVASLWFMSLFICIFATSILEMRWSGVG----IDDWWRNEQFW-VIGGVSSH 941
                  :|               : : | :|  | :  ||    |:  ||   ||  :   :
Db        431 IAASP-------------LALLAYAIPHM-FHAVGTASKINKGWR-YSFWSEVYETTMA 474

Qy        942 LFAVFQGLLKVIAGVDTSFTVTSKGGDDE----EFSELYTFKWTTLLIPPTTLLLLNFIG 997
               || |   :: :::     | || |||  |    :   :|         |   | |: | |
Db        475 LFLVRVTIVTLLSPSRGKFNVTDKGGLLEKGYFDLGAVY---------PNIILGLIMFGG 525

Qy        998 VVAGVSNAINNGYE-SWGPL-----FGKLFFAFWVIVHLYPFLKGL-VGRQ-------NR 1043
                 :  ||        || |:| |         |  : | :: |   |  : |||:       :|
Db        526 LARGV-------YELSFGHLDQIAERAYLLNSAWAMLSLIIILAAIAVGRETQQKRNSHR 578

Qy       1044 TPTIVIV 1050
                |  : |
Db        579 IPATIPV 585


RESULT 13
BCA3_ACEXY
ID   BCA3_ACEXY     STANDARD;      PRT;   745 AA.
AC   Q9WX61;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cellulose synthase 1 catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   BCSAI.
OS   Acetobacter xylinus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconacetobacter.
OX   NCBI_TaxID=28448;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 7664 / IFO 13693;
RX   MEDLINE=99310341; PubMed=10382968;
RA   Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA   Inoue Y.;
RT   "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT   7664: implication of a novel set of cellulose synthase genes.";
RL   DNA Res. 6:109-115(1999).

CC   -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC       uridine 5'-diphosphate glucose to cellulose. The thick cellulosic
CC       mats generated by this enzyme probably provide a specialized
CC       protective environment to the bacterium (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC       + {(1,4)-beta-D-glucosyl}(N+1).
CC   -!- COFACTOR: Magnesium (By similarity).
CC   -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC       (c-di-GMP) (By similarity).
CC   -!- PATHWAY: Bacterial cellulose biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- DOMAIN: There are two conserved domains in the globular part of
CC       the catalytic subunit: the N-terminal domain (domain A) contains
CC       the conserved DXD motif and is possibly involved in catalysis and
CC       substrate binding. The C-terminal domain (domain B) contains the
CC       QXXRW motif and is present only in processive glycosyl
CC       transferases. It could be involved in the processivity function of
CC       the enzyme, possibly required for holding the growing glycan chain
CC       in the active site.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB015802; BAA77585.1; -.
DR   InterPro; IPR003919; Cellsynth_A.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
DR   PRINTS; PR01439; CELLSNTHASEA.
KW   Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW   Transmembrane; Inner membrane.
FT   DOMAIN      147    240       CATALYTIC SUBDOMAIN A.
FT   DOMAIN      317    377       CATALYTIC SUBDOMAIN B.
FT   TRANSMEM     29     49       POTENTIAL.
FT   TRANSMEM    106    126       POTENTIAL.
FT   TRANSMEM    153    173       POTENTIAL.
FT   TRANSMEM    407    427       POTENTIAL.
FT   TRANSMEM    430    450       POTENTIAL.
FT   TRANSMEM    468    488       POTENTIAL.
FT   TRANSMEM    515    535       POTENTIAL.
FT   TRANSMEM    547    567       POTENTIAL.
FT   TRANSMEM    649    669       POTENTIAL.
FT   ACT_SITE    189    189       POTENTIAL.
FT   ACT_SITE    333    333       POTENTIAL.
FT   BINDING     236    236       SUBSTRATE (POTENTIAL).
FT   BINDING     238    238       SUBSTRATE (POTENTIAL).
SQ   SEQUENCE   745 AA;  83518 MW;  57EA0457A226F815 CRC64;

  Query Match             3.3%;  Score 194;  DB 1;  Length 745;
  Best Local Similarity   18.7%;  Pred. No. 1.5e-06;
  Matches  156;  Conservative   88;  Mismatches  221;  Indels  370;  Gaps   36;

Qy        276 IIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQ---FPKWFP------ 326
                :|: || :  ||               | ::| :  : : ::| | :   |  |
Db         53 MIVALVCVAVFFIVGRRKSRRTQVFLEVLSALVSLRY-LTWRLTETLDFDTWTQGILGVT 111

Qy        327 -IERETY-LDRLSLRFDKEGQPSQLAP-----------IDFFVSTVDPLKEPPLVTTNT 372
                :  | | |  | | : :   |   ||           :| |: | |   |   :   |
Db        112 LLLAELYALYMLFLSYFQTISPLHRAPLPLPANPDEWPTVDIFIPTYD---EALSIVRLT 168

Qy        373 VLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQ 432
                ||  | :|:| |||: |: |||
Db        169 VLGALGIDWPPDKVNVYILDDG---------------------------------- 190

Qy        433 QKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNV 492
                                 ||      |||     |  |:
Db        191 -----------------RR------EEFARFAEACGAR---------------------- 205
```

```
Qy        493 RDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNA 552
                                       |::|      | | ||| :|  ::     |
Db        206 ----------------------------YIARPD----NAHAKAGNLNYAIK----HTTG 229

Qy        553 PYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFF 612
               ::| ||||| |     :: :| :|:        :  :| |   |    |
Db        230 DHILILDCDH-IPTRAFLQISMGWMVS---DSNIALLQTPHHFYSPD------------ 272

Qy        613 DINMKGLDGIQGPIYVGTGCVPRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRK 672
                                  |:|    ||     : ||                  ||
Db        273 --------------------PFQRNLAVGY-----RTPPE---------------GN-- 289

Qy        673 QKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSS 732
                        ||:                          |::          |  ::
Db        290 -------------LFY--------------------------GVIQD---GNDFWDAT 305

Qy        733 VFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFK 792
               |                | | | ::||  |                  |:   :||||  |  :
Db        306 FF-------------CGSCAILRRKAIEEIG------------GFATETVTEDAHTALR 339

Qy        793 MHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGL 851
               |   || : |  ||      | |    |   : | :||| | |:||  ::  |    | ||
Db        340 MQRKGWSTAYLRIP---LASGLATERLITHIGQRMRWARGMIQIFRVDNPML----GSGL 392

Qy        852 KFLERFSYINSIVYPWTSIPLLAYCTLP----------AICLLTGKFITPELNNVASL 899
               |  :|  |::::   : :|| : :   |           |  |  | :  | :
Db        393 KLGQRLCYLSAMTSFFFAIPRVIFLASPLAFLFFSQNIIAASPLAVGVYAIPHM------ 446

Qy        900 WFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDT- 958
               | |:     ||:          ::  ||   ||      | ::     |  |   : |
Db        447 -FHSI-----ATA--------AKVNKGWR-YSFW------SEVYETVMALFLVRVTIVTM 485

Qy        959 ------SFTVTSKGG--DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGY 1010
                     | || |||  : |||    |:        |     ::  :|:: |
Db        486 LFPSKGKFNVTEKGGVLEREEFDLTATY-------PNIIFAIIMALGLLLRG--------- 529

Qy       1011 ESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWS----ILLASIFSL 1061
                    |:  :|    :|                |   :  :||    |:| :: |:
Db        530 -----LYALIFQHLDII-------------SERAYALNCIWSVISLIILMAVISV 566
```

RESULT 14
CBL_HUMAN
ID   CBL_HUMAN   STANDARD;   PRT;   906 AA.
AC   P22681;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   CBL E3 ubiquitin protein ligase (EC 6.3.2.-) (Signal transduction
DE   protein CBL) (Proto-oncogene c-CBL).
GN   CBL OR CBL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91232862; PubMed=2030914;
RA   Blake T.J., Shapiro M., Morse H.C. III, Langdon W.Y.;
RT   "The sequences of the human and mouse c-cbl proto-oncogenes show
RT   v-cbl was generated by a large truncation encompassing a proline-rich
RT   domain and a leucine zipper-like motif.";
RL   Oncogene 6:653-657(1991).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=99445925; PubMed=10514377;
RA   Joazeiro C.A., Wing S.S., Huang H.-K., Leverson J.D., Hunter T.,
RA   Liu Y.-C.;
RT   "The tyrosine kinase negative regulator c-Cbl as a RING-type,
RT   E2-dependent ubiquitin-protein ligase.";
RL   Science 286:309-312(1999).
RN   [3]
RP   INTERACTION WITH SLA AND ZAP70, AND MUTAGENESIS OF GLY-306.
RX   MEDLINE=99380595; PubMed=10449770;
RA   Tang J., Sawasdikosol S., Chang J.-H., Burakoff S.J.;
RT   "SLAP, a dimeric adapter protein, plays a functional role in T cell
RT   receptor signaling.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).
RN   [4]
RP   INTERACTION WITH SLA2.
RX   MEDLINE=21553259; PubMed=11696592;
RA   Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,
RA   Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,
RA   Yu S., Chan E., Wu X., Li C., Woisetschlager M., Aversa G.,
RA   Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,
RA   Mancebo H.S.Y., Wu J.;
RT   "Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel
RT   inhibitor of antigen receptor signaling.";
RL   J. Exp. Med. 194:1263-1276(2001).
RN   [5]
RP   INTERACTION WITH CD2AP.
RX   MEDLINE=21265017; PubMed=11067845;
RA   Kirsch K.H., Georgescu M.M., Shishido T., Langdon W.Y., Birge R.B.,
RA   Hanafusa H.;
RT   "The adapter type protein CMS/CD2AP binds to the proto-oncogenic
RT   protein c-Cbl through a tyrosine phosphorylation-regulated Src
RT   homology 3 domain interaction.";
RL   J. Biol. Chem. 276:4957-4963(2001).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 47-350.
RX   MEDLINE=99176421; PubMed=10078535;
RA   Meng W., Sawasdikosol S., Burakoff S.J., Eck M.J.;
RT   "Structure of the amino-terminal domain of Cbl complexed to its
RT   binding site on ZAP-70 kinase.";
RL   Nature 398:84-90(1999).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 47-434 IN COMPLEX WITH ZAP70
RP   AND UBE2L3.
RX   MEDLINE=20419298; PubMed=10966114;
RA   Zheng N., Wang P., Jeffrey P.D., Pavletich N.P.;
RT   "Structure of a c-Cbl-UbcH7 complex: RING domain function in
RT   ubiquitin-protein ligases.";
RL   Cell 102:533-539(2000).
CC   -!- FUNCTION: Participates in signal transduction in hematopoietic
CC   cells. Adapter protein that functions as a negative regulator of
CC   many signaling pathways that start from receptors at the cell
CC   surface. Acts as an E3 ubiquitin-protein ligase, which accepts
CC   ubiquitin from specific E2 ubiquitin-conjugating enzymes, and then
CC   transfers it to substrates promoting their degradation by the
CC   proteasome. Recognizes activated receptor tyrosine kinases,
CC   including PDGFA, EGF and CSF1, and terminates signaling.
CC   -!- SUBUNIT: Associates with NCK via its SH3 domain. The
CC   phosphorylated C-terminus interacts with CD2AP via its second SH3
CC   domain. Binds to ZAP70 and UBE2L3. Interacts with adapters SLA and
CC   SLA2.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- DOMAIN: The RING-type zinc finger domain mediates binding to an E2
CC   ubiquitin-conjugating enzyme.
CC   -!- PTM: Phosphorylated on tyrosine residues.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC   -!- SIMILARITY: Contains 1 UBA domain.
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC   WWW="http://www.infobiogen.fr/services/chromcancer/Genes/CBLID171.html".
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; X57110; CAA40393.1; -.
DR   PIR; A43817; A43817.
DR   PDB; 1B47; 27-APR-99.

```
DR   PDB; 1FBV; 17-JUL-00.
DR   PDB; 2CBL; 18-MAY-99.
DR   Genew; HGNC:1541; CBL.
DR   MIM; 165360; -.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0007048; P:oncogenesis; TAS.
DR   InterPro; IPR003153; Cbl_N.
DR   InterPro; IPR000449; UBA_domain.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02761; Cbl_N2; 1.
DR   Pfam; PF02762; Cbl_N3; 1.
DR   Pfam; PF02262; Cbl_N; 1.
DR   Pfam; PF00627; UBA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
DR   SMART; SM00165; UBA; 1.
DR   PROSITE; PS50030; UBA; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Ligase; Ubl conjugation pathway; Proto-oncogene; Zinc-finger;
KW   Phosphorylation; 3D-structure.
FT   ZN_FING      381    420       RING-TYPE.
FT   DOMAIN       357    476       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       477    688       PRO-RICH.
FT   DOMAIN       648    906       INTERACTS WITH CD2AP.
FT   DOMAIN       689    834       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       856    895       UBA.
FT   MOD_RES      700    700       PHOSPHORYLATION.
FT   MOD_RES      774    774       PHOSPHORYLATION.
FT   MUTAGEN      306    306       G->E: ABOLISHES INTERACTION WITH ZAP70,
FT                                 BUT DOES NOT AFFECT INTERACTION WITH SLA.
FT   HELIX         53     70
FT   TURN          71     71
FT   HELIX         73     75
FT   HELIX         84    101
FT   TURN         102    102
FT   HELIX        106    110
FT   TURN         111    111
FT   HELIX        113    136
FT   HELIX        137    141
FT   TURN         143    144
FT   HELIX        146    168
FT   HELIX        170    172
FT   HELIX        176    178
FT   HELIX        184    194
FT   TURN         195    196
FT   STRAND       199    201
FT   HELIX        202    212
FT   HELIX        218    228
FT   TURN         230    231
FT   STRAND       235    237
FT   HELIX        238    248
FT   HELIX        251    253
FT   HELIX        254    262
FT   TURN         263    263
FT   TURN         266    267
FT   STRAND       268    271
FT   HELIX        274    281
FT   HELIX        282    284
FT   TURN         285    286
FT   TURN         288    289
FT   STRAND       290    295
FT   TURN         297    298
FT   TURN         300    301
FT   STRAND       303    308
FT   TURN         310    311
FT   STRAND       314    317
FT   TURN         320    321
FT   HELIX        324    333
FT   TURN         334    335
FT   STRAND       339    340
FT   TURN         341    342
SQ   SEQUENCE   906 AA;  99646 MW;  7D686B050204AD8F CRC64;
```

```
  Query Match             2.4%;  Score 139.5;  DB 1;  Length 906;
  Best Local Similarity   19.4%;  Pred. No. 0.025;
  Matches  115;  Conservative   62;  Mismatches  194;  Indels  221;  Gaps   29;

Qy         37 QVCQICGDDVGLAPGGDPFVACNECAFPVCRDCY-EYERREGTQNCPQCKTRYKRLKGCQ 95
              |:|:|| ::       |  |    |   :|  |   ::  || | || |:     :|| :
Db        379 QLCKICAEN-------DKDVKIEPCGHLMCTSCLTSWQESEG-QGCPFCRC---EIKGTE 427

Qy         96 ---------RVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQA 146
                       | :|     | :   :  |:|  | :   :::    |    |       |
Db        428 PIVVDPFDPRGSGSLLRQGAEGAPSP-NYDDDDDERADDTLFM--MKELAGAKVERPPSP 484

Qy        147 FQLNPNVPLLTNGQMVDDIPPEQHALVP---SFMGGGGKRIHPLPYADPSLPVQPRSMD- 202
              | : |   |       :| :|  |   ||   | :|   |      : |  |||  |    |
Db        485 FSMAPQASLPPVPPRLDLLP--QRVCVPSSASALGTASKAASGSLHKDKPLPVPPTLRDL 542

Qy        203 --PSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSR 260
                |    |  |:        |:  ||:       |           | |  |:
Db        543 PPPPPPDRPYSVGA-------ESRPQRRPLPCTPG-----------DCPSRDKLP----- 579

Qy        261 KIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQ 320
                |:|||::                                              ||:
Db        580 --PVPSSRLGD-----------------------------------------------SWLPRP 594

Qy        321 FPKWFPIERETYLD---------RLSLRFDKEGQPSQLAP--------IDFFVSTVDPLK 363
                ||  |:   :  |         | ||| |     |||: |          | : |   :
Db        595 IPK-VPVSAPSSSDPWTGRELTNRHSLPF---SLPSQMEPRPDVPRLGSTFSLDTSMSMN 650

Qy        364 EPPLV------------TTNTVLSILSVDYPVDKV----SCYVSDDGAAM--------- 397
                |||             : | : |: :   || |:    |   :|   |
Db        651 SSPLVGPECDHPKIKPSSSANAIYSLAARPLPVPKLPPGEQCEGEEDTEYMTPSSRPLRP 710

Qy        398 ------------------LTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLK 439
                               |:||:                |||: :||     :   : :
Db        711 LDTSQSSRACDCDQQIDSCTYEAM---------------YNIQSQAPS--ITESSTFGE 752

Qy        440 DKVAANFV----RERRAMKREYEEFKVRINALVAK------AQKVPEEGWTMQDGTPWPG 489
                :||       |       |:  |  : |::|:      :      ||   ||   |
Db        753 GNLAAAHANTGPEESENEDDGYDVPKPPVPAVLARRTLSDISNASSSFGWLSLDGD--PT 810

Qy        490 NNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP---GYNHHKKAGA 538
                ||                  ||:::|      |  :|     |  :|||:
Db        811 TNV-----------------TEGSQVP-----ERPPKPFPRRINSERKAGS 839


RESULT 15
GLNA_SYNP2
ID   GLNA_SYNP2     STANDARD;      PRT;   473 AA.
AC   P28605;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
GN   GLNA.
OS   Synechococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32049;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PR-6;
RX   MEDLINE=93139025; PubMed=7678591;
RA   Wagner S.J., Thomas S.P., Kaufman R.I., Nixon B.T., Stevens S.E. Jr.;
RT   "The glnA gene of the cyanobacterium Agmenellum quadruplicatum PR-6
RT   is nonessential for ammonium assimilation.";
RL   J. Bacteriol. 175:604-612(1993).
CC   -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC       L-glutamine.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGONS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC   ---------------------------------------------------------------------
```

DR EMBL; Z13965; CAA78366.1; -.
DR HSSP; P06201; 1LGR.
DR InterPro; IPR001691; GLN_synth.
DR InterPro; IPR004809; GlnA.
DR InterPro; IPR001637; GlnA_adenyltn.
DR Pfam; PF00120; gln-synt; 1.
DR Pfam; PF03951; gln-synt_N; 1.
DR ProDom; PD001057; Gln_synt_C; 1.
DR TIGRFAMs; TIGR00653; GlnA; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase.
SQ SEQUENCE 473 AA; 53030 MW; 495C6C7A1EDBC576 CRC64;

Query Match 2.0%; Score 116; DB 1; Length 473;
Best Local Similarity 22.5%; Pred. No. 0.58;
Matches 75; Conservative 44; Mismatches 124; Indels 90; Gaps 19;

```
Qy        373 VLSILSVDYP-----------VDKVSCY--VSDDGAAMLTFEALSETS-----EFAKKW 413
              :: :  || |            :|: |    |  ||:::  ::|::|:      :    |
Db         19 IIDLKFVDLPGIWQHCSFYQDQLDEASFVDGVPFDGSSIRGWKAINESDMAMVPDPTTAW 78

Qy        414 V-PFCKRYNI-------EPRAPEWY-------FQQKIDYLKDKVAANFVRERRAMKREYE 458
              : ||||   :       |||  |||        |: :|||    | : : :      | |
Db         79 IDPFCKEKTLSLICSIKEPRTGEWYSRDPRSIAQKAVDYL----APSGIGDTAYFGPEAE 134

Qy        459 EF---KVRINALVAKAQKVPE--EG-WTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCE 512
               |    || :    |     :  || |    |   || |:   ||  | :          :
Db        135 FFVFDDVRFDQTENKGFYYVDSVEGRW--NSGRKEPGGNLAHKPGYKQGYFPVPPTDTLQ 192

Qy        513 GNELPRLVYVSREKRPGYNHHKK--AGAMNAL-VRVSAVLTNAPYLLNLDCDHYI--NNS 567
                    |: :::   |   || :   |  | |  | : :|  | ||:      |:  | :
Db        193 DMRTEMLLTMAKCGVPIEKHHHEVATGGQNELGFRFATLLKAADYLMTY---KYVIKNVA 249

Qy        568 KAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRH-------------DRYANRNVVFFDI 614
              :     : ||  ||          |     |:  |             |||||  : :
Db        250 RKYGRTVTFMPKPL---------FNDNGSGMHTHQSLWKEGQPLFWGDRYANLSQLALH- 299

Qy        615 NMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKT 647
                       |:| | :    ||  :  | |
Db        300 ------------YIG-GILKHAPALLAFSNPST 319
```

Search completed: December 15, 2003, 13:03:23
Job time : 26 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:00:29 ; Search time 45 Seconds
(without alignments)
6227.670 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5918	100.0	1086	10	Q9LLI3	Q9lli3 zea mays (m
2	5408.5	91.4	1059	10	Q9LLI4	Q9lli4 zea mays (m
3	5228	88.3	1094	10	Q9LLI2	Q9lli2 zea mays (m
4	4333.5	73.2	1091	10	Q93XQ1	Q93xq1 nicotiana a
5	4278.5	72.3	1084	10	O48947	O48947 arabidopsis
6	4262	72.0	1084	10	Q9FGF9	Q9fgf9 arabidopsis
7	4209.5	71.1	1088	10	Q9SJ22	Q9sj22 arabidopsis
8	4204.5	71.0	1069	10	Q9FIB9	Q9fib9 arabidopsis
9	4202	71.0	1081	10	O65338	O65338 arabidopsis
10	3984	67.3	1075	10	Q9LLI9	Q9lli9 zea mays (m
11	3977.5	67.2	1074	10	Q9LLI8	Q9lli8 zea mays (m
12	3967.5	67.0	1026	10	Q9SWW6	Q9sww6 arabidopsis
13	3963.5	67.0	1026	10	Q9XHP6	Q9xhp6 arabidopsis
14	3955.5	66.8	1032	10	Q8GSW2	Q8gsw2 populus tre
15	3945.5	66.7	1081	10	O48946	O48946 arabidopsis
16	3929.5	66.4	1067	10	Q9XGX6	Q9xgx6 gossypium h
17	3911	66.1	1076	10	Q9LLI5	Q9lli5 zea mays (m
18	3905.5	66.0	1077	10	Q9LLI6	Q9lli6 zea mays (m
19	3898.5	65.9	1065	10	Q9FHK6	Q9fhk6 arabidopsis
20	3886	65.7	1042	10	O81649	O81649 populus x c
21	3880.5	65.6	1079	10	Q9LLI1	Q9lli1 zea mays (m
22	3875.5	65.5	1065	10	O48948	O48948 arabidopsis
23	3835.5	64.8	1063	10	Q9AV71	Q9av71 oryza sativ
24	3826	64.7	1129	10	Q8LK26	Q8lk26 mesotaenium
25	3810	64.4	1055	10	Q8GZN8	Q8gzn8 arabidopsis
26	3804.5	64.3	1043	10	Q9FNC3	Q9fnc3 arabidopsis
27	3782	63.9	1065	10	Q9SKJ5	Q9skj5 arabidopsis
28	3476.5	58.7	978	10	Q9AXK0	Q9axk0 zinnia eleg
29	3469.5	58.6	974	10	P93155	P93155 gossypium h
30	3421	57.8	985	10	Q9C5Z8	Q9c5z8 arabidopsis
31	3420	57.8	985	10	Q8LPK5	Q8lpk5 arabidopsis
32	3410.5	57.6	974	10	Q8W1W0	Q8w1w0 gossypium h
33	3332.5	56.3	978	10	O81368	O81368 populus tre
34	3329.5	56.3	958	10	Q9SN37	Q9sn37 arabidopsis
35	3318	56.1	821	10	Q9LLI7	Q9lli7 zea mays (m
36	3165	53.5	939	10	Q943H3	Q943h3 oryza sativ
37	2907	49.1	685	10	P93156	P93156 gossypium h
38	2330.5	39.4	1181	10	Q9SRW9	Q9srw9 arabidopsis
39	2326	39.3	1145	10	Q9M9M4	Q9m9m4 arabidopsis
40	2324.5	39.3	1104	10	Q8GUZ9	Q8guz9 populus tre
41	2321.5	39.2	1170	10	Q9LHZ7	Q9lhz7 oryza sativ
42	2311	39.1	1145	10	Q9LFL0	Q9lfl0 arabidopsis
43	2293.5	38.8	1127	10	Q8W3F9	Q8w3f9 oryza sativ
44	2256	38.1	1111	10	Q9SZL9	Q9szl9 arabidopsis
45	2253	38.1	1127	10	Q93XQ0	Q93xq0 nicotiana a

ALIGNMENTS

RESULT 1
Q9LLI3
ID Q9LLI3 PRELIMINARY; PRT; 1086 AA.
AC Q9LLI3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cellulose synthase-7.
GN CESA-7.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene family.";
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200531; AAF89967.1; -.
DR InterPro; IPR005150; Cellulose_synt.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1086 AA; 122608 MW; DD03C73ABD13E2EB CRC64;

Query Match 100.0%; Score 5918; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
```

```
Db        61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy       121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Db       121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy       181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Db       181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy       241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHPVNDAFA 300

Db       241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy       301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Db       301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy       361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Db       361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy       421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Db       421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy       481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Db       481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy       541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Db       541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy       601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Db       601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy       661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Db       661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy       721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Db       721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy       781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Db       781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy       841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Db       841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy       901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Db       901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy       961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Db       961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy      1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Db      1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy      1081 CGLDCN 1086

Db      1081 CGLDCN 1086
```

RESULT 2

```
Q9LLI4
ID   Q9LLI4      PRELIMINARY;      PRT;  1059 AA.
AC   Q9LLI4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase-6.
GN   CESA-6.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20398328; PubMed=10938350;
RA   Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA   Xoconostle-Cazares B., Delmer D.P.;
RT   "A comparative analysis of the plant cellulose synthase (CesA) gene
RT   family.";
RL   Plant Physiol. 123:1313-1324(2000).
DR   EMBL; AF200530; AAF89966.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1059 AA;  120044 MW;  289DA26B25232249 CRC64;

  Query Match             91.4%;  Score 5408.5;  DB 10;  Length 1059;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  995;  Conservative   22;  Mismatches   36;  Indels    5;  Gaps    3;

Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLK 92

Db          3 QRNGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYEYERREGTQNCPQCKTRFKRFK 62

Qy         93 GCQRVTGDEEEDGVDDLDNEFNW-DGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNP 151

Db         63 GCARVPGDEEEDGVDDLENEFNWSDKHDSQYLAESMLHAHMSYGRGADLDGVPQPFHPIP 122

Qy        152 NVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYG 211

Db        123 NVPLLTNGQMVDDIPPDQHALVPSFVGGGGKRIHPLPYADPNLPVQPRSMDPSKDLAAYG 182

Qy        212 YGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINP 271

Db        183 YGSVAWKERMESWKQKQERMHQTRNDGGGDDGDDADLPLMDEARQPLSRKIPLPSSQINP 242

Qy        272 YRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 331

Db        243 YRMIIIIRLVVLCFFFHYRVMHPVPDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 302

Qy        332 YLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVS 391

Db        303 YLDRLSLRFDKEGHPSQLAPVDFFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVSCYVS 362

Qy        392 DDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERR 451

Db        363 DDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQQKIDYLKDKVAPNFVRERR 422

Qy        452 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDC 511

Db        423 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDV 482

Qy        512 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 571

Db        483 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 542

Qy        572 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 631

Db        543 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 602
```

```
Qy       632 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQK---KTTKPKTEKKKLLF 688
Db       603 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCICCCCFGNRKTKKKTKTSKPKFEKIKKL- 661

Qy       689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748
Db       662 FKKKENQAPAYALGEIDEAAPGAENEKASIVNQQKLEKKFGQSSVFVASTLLENGGTLKS 721

Qy       749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
Db       722 ASPASLLKEAIHVISCGYEDKTGWGKDIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRA 781

Qy       809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
Db       782 AFKGSAPLNLSDRFHQVLRWALGSIEILFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 841

Qy       869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 928
Db       842 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 901

Qy       929 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 988
Db       902 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 961

Qy       989 TLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1048
Db       962 TLLLLNFIGVVAGISNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1021

Qy      1049 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1086
Db      1022 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1059
```

```
RESULT 3
Q9LLI2
ID   Q9LLI2      PRELIMINARY;      PRT;  1094 AA.
AC   Q9LLI2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase-8.
GN   CESA-8.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20398328; PubMed=10938350;
RA   Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA   Xoconostle-Cazares B., Delmer D.P.;
RT   "A comparative analysis of the plant cellulose synthase (CesA) gene
RT   family.";
RL   Plant Physiol. 123:1313-1324(2000).
DR   EMBL; AF200532; AAF89968.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1094 AA;  122575 MW;  DFBBC18F49D23F5E CRC64;

  Query Match             88.3%;  Score 5228;  DB 10;  Length 1094;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy         1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57
Db         1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy        58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116
```

```
Db        60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy       117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172
Db       120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy       173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229
Db       177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy       230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFH 288
Db       237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy       289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348
Db       297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy       349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408
Db       357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy       409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468
Db       417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy       469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528
Db       477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy       529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588
Db       537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596

Qy       589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648
Db       597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy       649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
Db       657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy       709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
Db       717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy       769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
Db       777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy       829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
Db       837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy       889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
Db       897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy       949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
Db       957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy      1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
Db      1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy      1069 FLAKDDGPLLEECGLDCN 1086
Db      1077 FLAKSNGPLLEECGLDCN 1094
```

```
RESULT 4
Q93XQ1
```

```
ID   Q93XQ1      PRELIMINARY;      PRT;  1091 AA.
AC   Q93XQ1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cellulose synthase catalytic subunit.
GN   CESA1.
CS   Nicotiana alata (Winged tobacco) (Persian tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4087;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pistil;
RX   MEDLINE=21196092; PubMed=11299383;
RA   Doblin M.S., De Melis L., Newbigin E., Bacic A., Read S.M.;
RT   "Pollen Tubes of Nicotiana alata Express Two Genes from Different
RT   beta-Glucan Synthase Families.";
RL   Plant Physiol. 125:2040-2052(2001).
DR   EMBL; AF304374; AAK49454.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1091 AA;  122682 MW;  3FB8DF94B19D519B CRC64;

  Query Match             73.2%;  Score 4333.5;  DB 10;  Length 1091;
  Best Local Similarity   71.9%;  Pred. No. 0;
  Matches  794;  Conservative  123;  Mismatches  153;  Indels   35;  Gaps   12;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |:    |||||||||| |||  |        :| :||:|||||:: :   |:||:|||
Db          1 MDTKGRLVAGSHNRNEFVVINADDVGRVTSVKELSGQICQICGDEIEVTVDGEPFIACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHD- 119
              ||||||| ||||||||| | |||||||:||:||   || ||:|:|  ||||:||:: |:
Db         61 CAFPVCRQCYEYERREGNQACPQCKTRFKRIKGSPRVDGDDEDDEFDDLDHEFDYHGNPR 120

Qy        120 --SQSVAESMLYGHMSYGRGGDPNGA---------PQAFQLNPNVPLLTNGQMVDDIPPE 168
                |::   | |      ||| : | :          | |  ||  :|||| ||  | |  :
Db        121 YMSEAAFSSRL------GRGTNHNASGLTTPSEVDPAA--LNSEIPLLTYGQEDDTISAD 172

Qy        169 QHAL-VPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQR 227
              :||| :| || | ||::||:||:| |: : || ||| |||| ||||:||||||||:||::
Db        173 KHALIIPPFM-GRGKKVHPVPYSD-SMSLPPRPMDPKKDLAVYGYGTVAWKERMEDWKKK 230

Qy        228 QERMHQT----GNDGGGDDG---DDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRL 280
              |   |    |  |||:||   || ||| ||| || ||||:|: ||:::|||::|::||
Db        231 QNDKLQVVKHGGGKGGGNDGDELDDPDLPKMDEGRQPLSRKLPISSSRLSPYRLLILVRL 290

Qy        281 VVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRF 340
              |:| |||||: |||||:|||||||:|||||||:||| |||||||| ||||||||||
Db        291 AVVGLFFHYRITHPVNDAYALWLISIICEIWFAVSWIFDQFPKWFPIVRETYLDRLSLRY 350

Qy        341 DKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTF 400
              :|||:|| ||||| |||||||||||||:| |||||||:|||| ||||||||||||||||
Db        351 EKEGKPSGLAPIDIFVSTVDPLKEPPLITANTVLSILAVDYPEDKVSCYVSDDGAAMLTF 410

Qy        401 EALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEF 460
              ||||||||||:|||||||::|||||||||||| ||:||||:||  :||||||||||:||||
Db        411 EALSETSEFARKWVPFCKKFNIEPRAPEWYFSQKVDYLKNKVHPSFVRERRAMKRDYEEF 470

Qy        461 KVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLV 520
              ||||| ||| ||||||:|||||||||||||| ||||||||||||   |  | ||| ||||:
Db        471 KVRINGLVATAQKVPEDGWTMQDGTPWPGNLVRDHPGMIQVFLGNDGVRDIEGNVLPRLI 530

Qy        521 YVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDP 580
              |||||||||::||||||||||:|||||::|||||||:||||||||||::|||||||||
Db        531 YVSREKRPGFDHHKKAGAMNALMRVSAVISNAPYLLNVDCDHYINNSKALREAMCFMMDP 590

Qy        581 LLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALY 640
                                                                        Db        591 TSGKKICYVQFPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALY 650

Qy        641 GYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFPKKEENQSPAYA 700
              ||||||  ||| :|||||||  ||||||:||: |  |   : ||    | :|     :|
Db        651 GYDAPKKTKPPGKTCNCWPKW--CCCCFGSRKKHKKGKTTKDNKKKT--KTKEASPQIHA 706

Qy        701 LGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIH 760
              |  |:|   | ::||| :: | |||||||| ||| |||||:||    |: |||||||||
Db        707 LENIEEGIEGIDSEKATLMPQIKLEKKFGQSPVFVASTLLEDGGIPPGATSASLLKEAIH 766

Qy        761 VISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSD 820
              ||||||||||:||:|:||||||||||||||||||||:||:||| |||||||:||||
Db        767 VISCGYEDKTEWGREVGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRPAFKGSAPINLSD 826

Qy        821 RLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPA 880
              |||||||||||:||  | |||:|||| ||| |||||||||:||| ||:||:||| |||
Db        827 RLHQVLRWALGSVEILLSKHCPIWYGYGCGLKPLERFSYINSVVYPLTSLPLIAYCALPA 886

Qy        881 ICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSS 940
              :|||||||| ||::| ||: || ||| | |||:|||:| || |||||||||||||| ||
Db        887 VCLLTGKFIVPEISNYASILFMGLFIMIAATSVLEMQWGGVTIDDWWRNEQFWVIGGASS 946

Qy        941 HLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVA 1000
              ||||:|||||||:||| |||||||| || |||||| ||||:||||| |||::| |||:
Db        947 HLFALFQGLLKVLAGVSTSFTVTSKAADDGEFSELYLFKWTSLLIPPMTLLIINIIGVIV 1006

Qy       1001 GVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFS 1060
              |:|:||||||:|||||||:|||| ||||||||||::|||: |||::|||||||||||
Db       1007 GISDAINNGYDSWGPLFGRLFFALWVIVHLYPFLKGVMGRQNKVPTIIVVWSILLASIFS 1066

Qy       1061 LLWVRIDPFLAKDDGPLLEECGLDC 1085
              |||||::|| |:  | :|| |||||
Db       1067 LLWVRVNPFTAR-GGLVLEVCGLDC 1090
```

```
RESULT 5
O48947
ID   O48947      PRELIMINARY;      PRT;  1084 AA.
AC   O48947;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase catalytic subunit.
GN   H-A OR T22F8.250 OR AT4G39350.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98111412; PubMed=9445479;
RA   Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W.,
RA   Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
RA   Redmond J., Williamson R.E.;
RT   "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
RL   Science 279:717-720(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Zimmermann W., Grueneisen A., Wambutt R., Bancroft I.,
RA   Mewes H.W., Mayer K.F.X., Schueller C.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Zimmermann W., Grueneisen A., Wambutt R., Kalicki J., Wohldmann P.,
RA   Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
```

```
RN   [5]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF027173; AAC39335.1; -.
DR   EMBL; AL050351; CAB43650.1; -.
DR   EMBL; AL161595; CAB80598.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1084 AA;  122068 MW;  2F9B22D168D734E0 CRC64;

  Query Match             72.3%;  Score 4278.5;  DB 10;  Length 1084;
  Best Local Similarity   71.3%;  Pred. No. 0;
  Matches  782;  Conservative  127;  Mismatches  161;  Indels   27;  Gaps   12;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:||||||||  |:|   |      :  :| :|| ||||||:: |    : ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
              ||||||| |||||||||| | ||||||||||:||  || | ||||: :|||: ||:  | |
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEFD-HGMD 119

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
               :  ||: |   :: ||||  :  |        :||||      |:  ::|||:     |
Db        120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
              | |::| |: | | | | ||| | ||:| |||||||||:||| ||:|| |::    |:
Db        175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy        235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290
              ||:| |    |: || |:|:||| || ||||:|: ||:||||||:|: || :|| |||||
Db        235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy        291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350
              ::|||||: ||| |||||||||:|||||||||||||||||::|||:|| ||
Db        295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy        351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410
              |:| |||||||||||||:| |||||||:||||||||:|||||||||||||||||:|:|||
Db        355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSDTAEFA 414

Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470
              :|||||||::|||||||||| ||:||||:||   ||||||||||:||||||:||||||
Db        415 RKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530
              |||||||||||||||||||||||||||||||| ||  | :|||||||||||||||||:
Db        475 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590
              :|||||||||:|:||||||:||||||:||||||||||||:|:|||||||  ||||||||
Db        535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
              |||||||||||||:||||||||||||||||||||||||||||||||||:|||| |||
Db        595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKKP 654

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEA--A 708
              | :|||||||| | || | ||: ||   | : ||      :|     :|| :||
Db        655 PGKTCNCWPKWC-CLCC-GLRKKSKT---KAKDKKT---NTKETSKQIHALENVDEGVIV 706

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
              | :  ||     | |||||||| ||| | :|:|||  ::|||| ||:||| ||||||||
Db        707 PVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGYED 766

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
              ||:|||||||||||||||||||||||||:||:||| |||||||:||||||||||||
Db        767 KTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRW 826

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
              ||||:||| | |||:|||||||||:|||||||||:||: ||:||||:|||||||
Db        827 ALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTGKF 886

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
              | ||::| | : || :|| |  | ||||:| |||||||||||||||||| ||||||:|||
Db        887 IVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALFQG 946

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
              ||||:|||:|:|||||  ||  ||||| |||||||||||||::| |||: |||:||:|
Db        947 LLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAISN 1006

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
              ||:|||||||:|||| |||||||||||::|:|:: |||::|||||||| :|||||::|
Db       1007 GYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRVNP 1066

Qy       1069 FLAKDDGPLLEECGLDC 1085
              |:||  ||:|| |||:|
Db       1067 FVAK-GGPVLEICGLNC 1082


RESULT 6
Q9FGF9
ID   Q9FGF9      PRELIMINARY;      PRT;  1084 AA.
AC   Q9FGF9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase catalytic subunit.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB025637; BAB10307.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1084 AA;  122501 MW;  1520439A5053608C CRC64;

  Query Match             72.0%;  Score 4262;  DB 10;  Length 1084;
  Best Local Similarity   71.1%;  Pred. No. 0;
  Matches  778;  Conservative  126;  Mismatches  170;  Indels   20;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:|||||||  |:|   | :   :  :| :|| ||||  |:: |    |:|||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW--DGH 118
              ||||||| |||||||| | |||||||:|||||  || |||||| :|||||| :  :|
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDDIDDLDNEFEYGNNGI 120

Qy        119 DSQSVAESMLYGHMSYG-RGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL-VPSF 176
                 |:| |      : |     | : ||    |    :|||| |     :|  ::||| ||
Db        121 GFDQVSEGMSISRRNSGFPQSDLDSAPPGSQ----IPLLTYGDEDVEISSDRHALIVPPS 176

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERMHQTG 235
              :|| | |:||:  :||::    || | | |||| |||||||||:||| ||::| |::
Db        177 LGGHGNRVHPVSLSDPTVAAHPRPMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVR 236

Qy        236 NDGGGD--DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMH 293
              ::|  |  |||||| |:||| || |||||: ||:|||||:|::|||:|| |||||::|
Db        237 HEGDPDFEDGDDADFPMMDEGRQPLSRKIPIKSSKINPYRMLIVLRLVILGLFFHYRILH 296
```

```
Qy        294 PVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPID 353
              ||  ||:|||||||||||||||:||:|||||||:|||||||||||||||::|||:|| |:|:|
Db        297 PVKDAYALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVD 356

Qy        354 FFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKW 413
               |||||||||||||:| |||||||:||||||||:|||||||||||||||||||:|||:||
Db        357 VFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSETAEFARKW 416

Qy        414 VPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQK 473
              |||||:| |||||||||| |:||||:||  ||||||||||:||||||:|||||| |||
Db        417 VPFCKKYCIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQK 476

Qy        474 VPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHH 533
              |||:|||||||||||||:|||||||||||||  |  | | |||||||||||||||::||
Db        477 VPEDGWTMQDGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPGFDHH 536

Qy        534 KKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQ 593
              |||||||:|:||| ||:|||||||:||||||||||::|||||||| |||:|||||||
Db        537 KKAGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQ 596

Qy        594 RFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSR 653
              ||||||||||:|||||||||||||||:||||||||||||||||||:|||| || | :
Db        597 RFDGIDRHDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKGPRK 656

Qy        654 TCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDE--AAPGA 711
              |||||||||    |||:|| :|      :|||    |  |      :||  |:|     |:
Db        657 TCNCWPKWCL--LCFGSRKNRKAKTVAADKKK----KNREASKQIHALENIEEGRVTKGS 710

Qy        712 ENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTD 771
              |::    | |||||||| ||| |  :||||  ::|||| ||||||  ||||||||||:
Db        711 NVEQSTEAMQMKLEKKFGQSPVFVASARMENGGMARNASPACLLKEAIQVISCGYEDKTE 770

Qy        772 WGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALG 831
              |||||||||||||||||||||| |||||:|| ||  |||||||:||||||||||||||||
Db        771 WGKEIGWIYGSVTEDILTGFKMHSHGWRSVYCTPKLAAFKGSAPINLSDRLHQVLRWALG 830

Qy        832 SIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITP 891
              |:||| | |||:||||||||:||| |||||:|||||:||: ||:|||||||||||| |
Db        831 SVEIFLSRHCPIWYGYGGGLKWLERLSYINSVVYPWTSLPLIVYCSLPAICLLTGKFIVP 890

Qy        892 ELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLK 951
              |::| ||: ||:||  |  | ||||:|  ||||||||||||||||||:||||:||||||
Db        891 EISNYASILFMALFSSIAITGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLK 950

Qy        952 VIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYE 1011
              |:|||||:||||||  || |||:|| ||||:||||| |||::| |||: |||:||:|||:
Db        951 VLAGVDTNFTVTSKAADDGEFSDLYLFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYD 1010

Qy       1012 SWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLA 1071
              |||||||:|||| |||:|||||||||:|:|:| |||::|||||||| :|||||::||:|
Db       1011 SWGPLFGRLFFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVA 1070

Qy       1072 KDDGPLLEECGLDC 1085
              |  ||:|| |||||
Db       1071 K-GGPILEICGLDC 1083


RESULT 7
Q9SJ22
ID   Q9SJ22      PRELIMINARY;      PRT;  1088 AA.
AC   Q9SJ22;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative cellulose synthase catalytic subunit.
GN   AT2G21770.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007019; AAD20396.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
SQ   SEQUENCE   1088 AA;  123446 MW;  060571118600DC9F CRC64;

  Query Match             71.1%;  Score 4209.5;  DB 10;  Length 1088;
  Best Local Similarity   70.2%;  Pred. No. 0;
  Matches  769;  Conservative  127;  Mismatches  179;  Indels   21;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:||||||||| |:|  |     :   | :|| |:|| |:: |    |:||:||||
Db          1 MNTGGRLIAGSHNRNEFVLINADDTARIRSAEELSGQTCKICRDEIELTDNGEPPFIACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              |||| || |||||||||  | |||| |||||:||  || ||||:| :|||::||    | |
Db         61 CAFPTCRPCYEYERREGNQACPQCGTRYKRIKGSPRVEGDEEDDDIDDLEHEFY--GMDP 118

Qy        121 QSVAESML-YGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
              : | |: | |   ::  ||| |      :      ||||      |:  ::|||:     |
Db        119 EHVTEAALYYMRLNTGRGTDEVSHLYSASPGSEVPLLTYCDEDSDMYSDRHALIVPPSTG 178

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
               | |:| :|: |    :  | | | |||  |||||||||:||| ||::| |::    ::
Db        179 LGNRVHHVPFTDSFASIHTRPMVPQKDLTVYGYGSVAWKDRMEVWKKQQIEKLQVVKNER 238

Qy        235 GNDGGG-----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHY 289
              |||  |      |: ||  ||:||| || ||||:|: ||:|||||:|  || :|| ||||
Db        239 VNDGDGDGFIVDELDDPGLPMMDEGRQPLSRKLPIRSSRINPYRMLIFCRLAILGLFFHY 298

Qy        290 RVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQL 349
              |::|||||| ||| |||||||||:||||||||:|||||||||||||::|||:||:|
Db        299 RILHPVNDAFGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSEL 358

Qy        350 APIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEF 409
              ||:| |||||||||||||:| |||||||:|||||:||:|||||||||||||| |:||
Db        359 APVDVFVSTVDPLKEPPLITANTVLSILAVDYPVEKVACYVSDDGAAMLTFEALSYTAEF 418

Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469
              |:|||||:::||||||||| ||:|||| ||   || |||||||:||||||:|||||:
Db        419 ARKWVPFCKKFSIEPRAPEWYFSQKMDYLKHKVDPAFVMERRAMKRDYEEFKVKINALVS 478

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529
               :|||||:|||||||||||||||||||||||||| ||  | :|||||||||||||||||
Db        479 VSQKVPEDGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVCDMDGNELPRLVYVSREKRPG 538

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589
              ::|||||||||:|:|||||:||||||:|||||||||||:|||||||  |||:|||
Db        539 FDHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIREAMCFMMDPQSGKKICYV 598

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649
              |||||||||||||:|||||||||||||||||||||||||||||||||||:|||| |:
Db        599 QFPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKQ 658
```

```
Qy       650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAP 709
Db       659 PPGRTCNCWPKWC-CLCC-GMRK-KKTGKVKDNQRK----KPKETSKQIHALEHIEEGLQ 711

Qy       710 GAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDK 769
Db       712 VTNAENNSETAQLKLEKKFGQSPVLVASTLLLNGGVPSNVNPASLLRESIQVISCGYEEK 771

Qy       770 TDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWA 829
Db       772 TEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRWA 831

Qy       830 LGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI 889
Db       832 LGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLLVYCSLPAICLLTGKFI 891

Qy       890 TPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGL 949
Db       892 VPEISNYAGILFLLMFMSIAVTGILEMQWGKIGIDDWWRNEQFWVIGGVSSHLFALFQGL 951

Qy       950 LKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNG 1009
Db       952 LKVLAGVSTNFTVTSKAADDGEFSELYIFKWTSLLIPPTTLLIINIVGVIVGVSDAINNG 1011

Qy      1010 YESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPF 1069
Db      1012 YDSWGPLFGRLFFALWVIVHLYPFLKGLLGKQDRVPTIILVWSILLASILTLLWVRVNPF 1071

Qy      1070 LAKDDGPLLEECGLDC 1085
Db      1072 VSK-DGPVLEICGLDC 1086
```

```
RESULT 8
Q9FIB9
ID   Q9FIB9      PRELIMINARY;      PRT;  1069 AA.
AC   Q9FIB9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase catalytic subunit.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=99156233; PubMed=10048488;
RA   Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
RT   Sequence features of the regions of 1,081,958 bp covered by seventeen
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:379-391(1998).
DR   EMBL; AB016893; BAB09408.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   1069 AA;  120861 MW;  D03ED5C578DB3E7C CRC64;

  Query Match             71.0%;  Score 4204.5;  DB 10;  Length 1069;
  Best Local Similarity   70.6%;  Pred. No. 0;
  Matches  770;  Conservative  126;  Mismatches  166;  Indels   29;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVEELSGQTCQICGDEIELSVDGESFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
```

```
Db         61 CAFPVCRPCYEYERREGNQSCPQCKTRYKRIKGSPRVEGDEEDDGIDDLDFEFDY----S 116

Qy        121 QSVAESMLYG--HMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMG 178
Db        117 RSGLESETFSRRNSEFDLASAPPGS--------QIPLLTYGEEDVEISSDSHALIVSPSP 168

Qy        179 GGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQT---G 235
Db        169 GHIHRVHQPHFPDPA--AHPRPMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKYQVVKHD 226

Qy        236 NDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHPV 295
Db        227 GDSSLGDGDDADIPMMDEGRQPLSRKVPIKSSKINPYRMLIVLRLVILGLFFHYRILHPV 286

Qy        296 NDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFF 355
Db        287 NDAYALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSELAGVDVF 346

Qy        356 VSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVP 415
Db        347 VSTVDPMKEPPLITANTVLSILAVDYPVDRVACYVSDDGAAMLTFEALSETAEFARKWVP 406

Qy        416 FCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVP 475
Db        407 FCKKYTIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQKVP 466

Qy        476 EEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKK 535
Db        467 EEGWTMQDGTPWPGNNVRDHPGMIQVFLGNNGVRDVENNELPRLVYVSREKRPGFDHHKK 526

Qy        536 AGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRF 595
Db        527 AGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRF 586

Qy        596 DGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTC 655
Db        587 DGIDKSDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKTKRMTC 646

Qy        656 NCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAEN-E 714
Db        647 NCWPKWCLFCC--GLRKNRKS--KTTDKKK----KNREASKQIHALENIEEGTKGTNDAA 698

Qy        715 KAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGK 774
Db        699 KSPEAAQLKLEKKFGQSPVFVASAGMENGGLARNASPASLLREAIQVISCGYEDKTEWGK 758

Qy        775 EIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIE 834
Db        759 EIGWIYGSVTEDILTGFKMHSHGWRSVYCTPKIPAFKGSAPINLSDRLHQVLRWALGSVE 818

Qy        835 IFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELN 894
Db        819 IFLSRHCPIWYGYGGGLKWLERLSYINSVVYPWTSIPLLVYCSLPAICLLTGKFIVPEIS 878

Qy        895 NVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIA 954
Db        879 NYASILFMALFGSIAVTGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLA 938

Qy        955 GVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWG 1014
Db        939 GVETNFTVTSKAADDGEFSELYIFKWTSLLIPPTTLLIINVIGVIVGISDAISNGYDSWG 998

Qy       1015 PLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDD 1074
Db        999 PLFGRLFFAFWVILHLYPFLKGLLGKQDRMPTIILVWSILLASILTLLWVRVNPFVAK-G 1057

Qy       1075 GPLLEECGLDC 1085
Db       1058 GPILEICGLDC 1068
```

```
RESULT 9
```

```
O65338
ID   O65338      PRELIMINARY;      PRT;  1081 AA.
AC   O65338;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase (Fragment).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Wu L., Joshi C.P., Chiang V.L.;
RT   "AraxCelA, a new member of the cellulose synthase gene family from
RT   Arabidopsis (Accession No. AF062485) (PGR98-114).";
RL   Plant Physiol. 117:1125-1125(1998).
DR   EMBL; AF062485; AAC29067.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   1081 AA;  122446 MW;  DC59A35A1713FD9F CRC64;

  Query Match             71.0%;  Score 4202;  DB 10;  Length 1081;
  Best Local Similarity   70.8%;  Pred. No. 0;
  Matches  770;  Conservative  126;  Mismatches  172;  Indels   20;  Gaps   10;

Qy          7 LVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVC 66
              |:|||||||| |:|  | :   :  :| :|| |||| |:: |   |:|||||||||||||
Db          4 LIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNECAFPVC 63

Qy         67 RDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW--DGHDSQSVA 124
              | |||||||| | |||||||:|||||  || |||||| :|||||| :  :|     |:
Db         64 RPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDDIDDLDNEFEYGNNGIGFDQVS 123

Qy        125 ESMLYGHMSYG-RGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL-VPSFMGGGGK 182
              | |      : |    | : ||   |    :|||| |    :|  ::||| ||  :|| |
Db        124 EGMSISRRNSGFPQSDLDSAPPGSQ----IPLLTYGDEDVEISSDRHALIVPPSLGGHGN 179

Qy        183 RIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERMHQTGNDGGGD 241
              |:||:  :||::    | | | |||| |||||||||:||| ||::| |::    ::|  |
Db        180 RVHPVSLSDPTVAAHRRLMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVRHEGDPD 239

Qy        242 --DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAF 299
                |||||| |:||| || || ||| |:|:|||||:|::|||:|| |||||::||| ||:
Db        240 FEDGDDADFPMMDEGRQPLSMKIPIKSSKINPYRMLIVLRLVILGLFFHYRILHPVKDAY 299

Qy        300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTV 359
              |||||||||||||:||:|||||||:|||||||||||||::|||:|| |:|:| |||||
Db        300 ALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVDVFVSTV 359

Qy        360 DPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 419
              ||||||||:| |||||||:||||||||:||||||||||||||||||:|||:|||||||:
Db        360 DPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSETAEFARKWVPFCKK 419

Qy        420 YNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 479
              | |||||||||| |:||||:||   ||||||||||:||||||:|||||| |||||:||
Db        420 YCIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQKVPEDGW 479

Qy        480 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAM 539
              ||||||||||:|||||||||||||  |  | | |||||||||||||||::||||||||
Db        480 TMQDGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPGFDHHKKAGAM 539

Qy        540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 599
              |:|:||| ||:|||||||:||||||||||::|||||||  |||:|||||||||||||||
Db        540 NSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRFDGID 599

Qy        600 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 659
```

```
                ||||:|||||||||||||||:||||||||||||||||||:|||| || | :||||||
Db        600 RHDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKGPRKTCNCWP 659

Qy        660 KWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA--PGAENEKAG 717
              |||    |||:|| :|      :|||    | |     :||  |:|         |::
Db        660 KWCL--LCFGSRKNRKAKTVAADKKK----KNREASKQIHALENIEEGRGHKVLNVEQST 713

Qy        718 IVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIG 777
              | ||:||:||| ||| |  ||||| ::|||| |||||| ||| |||||:||||||
Db        714 EAMQMKLQKKYGQSPVFVASARLENGGMARNASPACLLKEAIQVISRGYEDKTEWGKEIG 773

Qy        778 WIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFF 837
              |||||||||||| ||| |||| :|| ||  |||||||:||||||||||||||:|||
Db        774 WIYGSVTEDILTGSKMHSHGWRHVYCTPKLAAFKGSAPINLSDRLHQVLRWALGSVEIFL 833

Qy        838 SNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVA 897
              | |||:|||||||||:||| |||||:||||||:||: ||:|||||||||||| ||::| |
Db        834 SRHCPIWYGYGGGLKWLERLSYINSVVYPWTSLPLIVYCSLPAICLLTGKFIVPEISNYA 893

Qy        898 SLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVD 957
              |: ||:||  |  | ||||:|  |||||||||||||||||:||||:|||||||:||||
Db        894 SILFMALFSSIAITGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLAGVD 953

Qy        958 TSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLF 1017
              |:||||||  || |||:|| ||||:||||| |||::| |||: |||:||:|||:||||||
Db        954 TNFTVTSKAADDGEFSDLYLFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYDSWGPLF 1013

Qy       1018 GKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPL 1077
              |:|||| |||:||||||||:|:|:| |||::|||||||| :|||||::||:||  ||:
Db       1014 GRLFFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVAK-GGPI 1072

Qy       1078 LEECGLDC 1085
              || |||||
Db       1073 LEICGLDC 1080


RESULT 10
Q9LLI9
ID   Q9LLI9      PRELIMINARY;      PRT;  1075 AA.
AC   Q9LLI9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cellulose synthase-1.
GN   CESA-1.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20398328; PubMed=10938350;
RA   Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA   Xoconostle-Cazares B., Delmer D.P.;
RT   "A comparative analysis of the plant cellulose synthase (CesA) gene
RT   family.";
RL   Plant Physiol. 123:1313-1324(2000).
DR   EMBL; AF200525; AAF89961.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
SQ   SEQUENCE   1075 AA;  121181 MW;  67A4FBD97A811F33 CRC64;

  Query Match             67.3%;  Score 3984;  DB 10;  Length 1075;
  Best Local Similarity   66.4%;  Pred. No. 0;
  Matches  738;  Conservative  143;  Mismatches  167;  Indels   64;  Gaps   15;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD-PG-PKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
              | |: |:|||||||| |:|| ||| ||  ||  :  |||||||||| ||::  || ||||
Db          1 MAANKGMVAGSHNRNEFVMIRHDGDVPGSAKPTKSANGQVCQICGDSVGVSATGDVFVAC 60
```

```
Qy        59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
             ||||||||| ||||||:|| | |||||||||| ||  || |||:|: |||||||||:
Db        61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGDEDEEDVDDLDNEFNY--- 117

Qy       119 DSQSVAESMLYGHMSYGRG------GDPNGAPQAFQLNPN--VPLLTNGQMVD----DIP 166
                           |:|      ||      : :  |:  :| ||:|| :     |
Db       118 ------------KQGSGKGPEWQLQGDDADLSSSARHEPHHRIPRLTSGQQISGEIPDAS 165

Qy       167 PEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQ 226
             |::|::          |     | |||:||  | :|||||| :||  || ||||:|:|:
Db       166 PDRHSI----------RSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRV 215

Qy       227 RQER-MHQTGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMI 275
             :|:: | |  |          :| | :|:   : ::|:|| ||| :|: |:|:| ||::
Db       216 KQDKNMMQVTNKYPEARGGDMEGTGSNGE--XMQMVDDARLPLSRIVPISSNQLNLYRVV 273

Qy       276 IIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDR 335
             ||:||::| ||| ||| ||| ||: |||:|||||:|||:||:|||||||:|| |||||||
Db       274 IILRLIILCFFFQYRVSHPVRDAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDR 333

Qy       336 LSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGA 395
             |:||:|:||:|||||||| |||||||||||||:| |||||||||||||||||||||||:
Db       334 LALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILSVDYPVDKVSCYVSDDGS 393

Qy       396 AMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKR 455
             ||||||:||||:|||:|||||||::|||||||:|| |||||||||:  :||:|||||||
Db       394 AMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKR 453

Qy       456 EYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNE 515
             ||||||||||||||||||||||||| ||| ||||| |||||||||||| ||||| :|||
Db       454 EYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNE 513

Qy       516 LPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMC 575
             |||||||||||||: |||||||||:|||||||  ||||:|||||  |:|||::||||
Db       514 LPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKALREAMC 573

Qy       576 FMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFR 635
             ||||| ||:| |||||||||||||| ||||||||:|||||||||||||||||:|||||| |
Db       574 FMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFN 633

Qy       636 RQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQ 695
             ||||||||      |     :  |      || | ||:|  :   :: :     |: |:
Db       634 RQALYGYD------PVLTEADLEPNIVIKSCC-GRRKKKNKSYMDSQSR---IMKRTESS 683

Qy       696 SPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755
             :| : : :|:|    | |:|:: :::|:||||:|||| :|: || :  ||    | :|||||
Db       684 APIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPASLL 743

Qy       756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815
             ||||||||||||||:|||||||||||||||||||||  ||:||||:| |  ||||||
Db       744 KEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFKGSAP 803

Qy       816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875
             :|||||:|||||||||:||   | |||:|||| | || ||| :|||:|||| |||||:||
Db       804 INLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITSIPLIAY 863

Qy       876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935
             | |||||||| ||| ||::| | ::|: ||  |||| |||:|||||||:||||||||||
Db       864 CVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVI 923

Qy       936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLN 994
             || |:|||||||||||:||:||:|||||   | | :|:||| ||||:|||||||:|:::|
Db       924 GGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTVLVIN 983

Qy       995 FIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSIL 1054
              :|:|||:| |||:||:|||||||||||: |||:|||||||||:||||||||||||||||
Db       984 LVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQNRTPTIVIVWSIL 1043

Qy      1055 LASIFSLLWVRIDPFLA-KDDGPLLEECGLDC 1085
             ||||||||||:||||::        | :||::|
Db      1044 LASIFSLLWVKIDPFISPTQKAAALGQCGVNC 1075
```

```
RESULT 11
Q9LLI8
ID   Q9LLI8      PRELIMINARY;      PRT;  1074 AA.
AC   Q9LLI8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cellulose synthase-2.
GN   CESA-2.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20398328; PubMed=10938350;
RA   Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA   Xoconostle-Cazares B., Delmer D.P.;
RT   "A comparative analysis of the plant cellulose synthase (CesA) gene
RT   family.";
RL   Plant Physiol. 123:1313-1324(2000).
DR   EMBL; AF200526; AAF89962.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
DR   SMART; SM00184; RING; 1.
SQ   SEQUENCE   1074 AA;  121187 MW;  A116E66F0564E210 CRC64;

  Query Match             67.2%;  Score 3977.5;  DB 10;  Length 1074;
  Best Local Similarity   66.1%;  Pred. No. 0;
  Matches  734;  Conservative  149;  Mismatches  165;  Indels   63;  Gaps   15;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD-PGP-KPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
              | |: |:|||||||||| |:|| ||| | | || :  |||||||||| ||::  || ||||
Db          1 MAANKGMVAGSHNRNEFVMIRHDGDAPVPAKPTKSANGQVCQICGDTVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
             |||||||||| |||||:|| | |||||||||| ||  || |||:|: |||||||||:
Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGDDEEEDVDDLDNEFNY--- 117

Qy        119 DSQSVAESMLYGHMSYGRG------GDPNGAPQAFQLNPN--VPLLTNGQMVD----DIP 166
                            |:|      ||      : : :|:  :| ||:|| :     |
Db        118 ------------KQGNGKGPEWQLQGDDADLSSSARHDPHHRIPRLTSGQQISGEIPDAS 165

Qy        167 PEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQ 226
             |::|::          |     | |||:||  | :|||||| :||  || ||||:|:|:
Db        166 PDRHSI----------RSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRV 215

Qy        227 RQER-MHQTGN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMII 276
             :|:: | |  |          :| | :|:  |: ::|:||  ||| :|: |:|:| ||::|
Db        216 KQDKNMLQVTNKYPEARGDMEGTGSNGE--DMQMVDDARLPLSRIVPISSNQLNLYRIVI 273

Qy        277 IIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRL 336
             |:|||::| ||| ||: ||| :|: |||:||||:|||:||:|||||||:|| |||||||
Db        274 ILRLIILCFFFQYRISHPVRNAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDRL 333

Qy        337 SLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAA 396
             :||:|:||:|||||||| ||||||||||||:| |||||||:|||||||||||||||:|
Db        334 ALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVSCYVSDDGSA 393

Qy        397 MLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKRE 456
             |||||:||||:|||:|||||||::|||||||:|| |||||||||:  :||:|||||||||
Db        394 MLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKRE 453

Qy        457 YEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNEL 516
             |||||:|||||||||||||||||| ||| ||||| |||||||||||| ||||| :||||
Db        454 YEEFKIRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNEL 513

Qy        517 PRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCF 576
             |||||||||||||: |||||||||:|||||||  ||||:|||||  |:|||::|||||
Db        514 PRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKALREAMCF 573
```

```
Qy       577 MMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRR 636
Db       574 MMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFNR 633

Qy       637 QALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQS 696
Db       634 QALYGYD------PVLTEADLEPNIVVKSCC-GRRKRKNKSYMDSQSR---IMKRTESSA 683

Qy       697 PAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLK 756
Db       684 PIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPASLLK 743

Qy       757 EAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPL 816
Db       744 EAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFKGSAPI 803

Qy       817 NLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYC 876
Db       804 NLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITSVPLIAYC 863

Qy       877 TLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIG 936
Db       864 VLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVIG 923

Qy       937 GVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNF 995
Db       924 GTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTVLVINL 983

Qy       996 IGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILL 1055
Db       984 VGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQNRTPTIVIVWSILL 1043

Qy      1056 ASIFSLLWVRIDPFLA-KDDGPLLEECGLDC 1085
Db      1044 ASIFSLLWVKIDPFISPTQKAAALGQCGVNC 1074
```

RESULT 12
Q9SWW6
ID Q9SWW6 PRELIMINARY; PRT; 1026 AA.
AC Q9SWW6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cellulose synthase catalytic subunit (AT5g17420/T10B6_80).
GN IRX3 OR T10B6_80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. LANDSBERG ERECTA;
RX MEDLINE=99264300; PubMed=10330464;
RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
RT synthase required for secondary cell wall synthesis.";
RL Plant Cell 11:769-780(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF091713; AAD40885.1; -.
DR EMBL; AL391142; CAC01737.1; -.
DR EMBL; AY139754; AAM98075.1; -.
DR InterPro; IPR005150; Cellulose_synt.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1026 AA; 115797 MW; 503BFBC78BE6E511 CRC64;

Query Match 67.0%; Score 3967.5; DB 10; Length 1026;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 736; Conservative 129; Mismatches 145; Indels 91; Gaps 16;

```
Qy         1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Db         1 MEASAGLVAGSHNRNELVVIHNHEE--PKPLKNLDGQFCEICGDQIGLTVEGDLFVACNE 58

Qy        61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
Db        59 CGFPACRPCYEYERREGTQNCPQCKTRYKRLRGSPRVEGDEDEEDIDDIEYEFNIE-HEQ 117

Qy       121 ---QSVAESMLYGHMSYGRG--GDPNGA-PQAFQLNPNVPLLTNGQMVDDIPPEQHALVP 174
Db       118 DKHKHSAEAMLYGKMSYGRGPEDDENGRFP---------PVIAGGHSGE---------- 157

Qy       175 SFMGGGG---------KRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWK 225
Db       158 -FPVGGGYGNGEHGLHKRVHPY--------------PSSEAGSEG----GWRERMDDWK 197

Qy       226 QRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGF 285
Db       198 L------QHGNLGPEPD-DDPEMGLIDEARQPLSRKVPIASSKINPYRMVIVARLVILAV 250

Qy       286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQ 345
Db       251 FLRYRLLNPVHDALGLWLTSVICEIWFAVSWILDQFPKWFPIERETYLDRLSLRYEREGE 310

Qy       346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405
Db       311 PNMLAPVDVFVSTVDPLKEPPLVTSNTVLSILAMDYPVEKISCYVSDDGASMLTFESLSE 370

Qy       406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465
Db       371 TAEFARKWVPFCKKFSIEPRAPEMYFTLKVDYLQDKVHPTFVKERRAMKREYEEFKVRIN 430

Qy       466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525
Db       431 AQVAKASKVPLEGWIMQDGTPWPGNNTKDHPGMIQVFLGHSGGFDVEGHELPRLVYVSRE 490

Qy       526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585
Db       491 KRPGFQHHKKAGAMNALVRVAGVLTNAPFMLNLDCDHYVNNSKAVREAMCFLMDPQIGKK 550

Qy       586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645
Db       551 VCYVQFPQRFDGIDTNDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFKRQALYGYEPP 610

Qy       646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705
Db       611 KGPKRPKMIS------CGCCPCFGRRRKNKK------------FSKNDMNGDVAALG--- 649

Qy       706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765
Db       650 ----GAEGDKEHLMSEMNFEKTFGQSSIFVTSTLMEEGGVPPSSSPAVLLKEAIHVISCG 705

Qy       766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825
```

```
Db        706 YEDKTEWGTELGWIYGSITEDILTGFKMHCRGWRSIYCMPKRPAFKGSAPINLSDRLNQV 765

Qy        826 LRWALGSIEIFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLL 884

Db        766 LRWALGSVEIFFSRHSPLWYGYKGGKLKWLERFAYANTTIYPFTSIPLLAYCILPAICLL 825

Qy        885 TGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFA 944

Db        826 TDKFIMPPISTFASLFFISLFMSIIVTGILELRWSGVSIEEWWRNEQFWVIGGISAHLFA 885

Qy        945 VFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004

Db        886 VVQGLLKILAGIDTNFTVTSKATDDDDFGELYAFKWTTLLIPPTTVLIINIVGVVAGISD 945

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064

Db        946 AINNGYQSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPTIVVIWSVLLASIFSLLWV 1005

Qy       1065 RIDPFLAKDDGPLLEECGLDC 1085

Db       1006 RIDPFVLKTKGPDTSKCGINC 1026
```

RESULT 13
Q9XHP6
ID Q9XHP6 PRELIMINARY; PRT; 1026 AA.
AC Q9XHP6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN IRX3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RX MEDLINE=99264300; PubMed=10330464;
RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
RT synthase required for secondary cell wall synthesis.";
RL Plant Cell 11:769-780(1999).
DR EMBL; AF088917; AAD32031.1; -.
DR InterPro; IPR005150; Cellulose_synt.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1026 AA; 115858 MW; 453BFD1D283C4D70 CRC64;

Query Match 67.0%; Score 3963.5; DB 10; Length 1026;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 736; Conservative 128; Mismatches 146; Indels 91; Gaps 16;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db          1 MEASAGLVAGSHNRNELVVIHNHEE--PKPLKNLDGQFCEICGDQIGLTVEGDLFVACNE 58

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db         59 CGFPACRPCYEYERREGTQNCPQCKTRYKRLRGSPRVEGDEDEEDIDDIEYEFNIE-HEQ 117

Qy        121 ---QSVAESMLYGHMSYGRG--GDPNGA-PQAFQLNPNVPLLTNGQMVDDIPPEQHALVP 174

Db        118 DKHKHSAEAMLYGKMSYGRGPEDDENGRFP---------PVIAGGHSGE----------- 157

Qy        175 SFMGGGG---------KRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWK 225

Db        158 -FPVGGGYGNGEHGLHKRVHPY--------------PSSEAGSEG----GWRERMDDWK 197
```

```
Qy        226 QRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285

Db        198 L------QHGNLGPEPD-DDPEMGLIDEARQPLSRKVPIASSKINPYRMVIVARLVILAV 250

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQ 345

Db        251 FLRYRLLNPVHDALGLWLTSVICEIWFAVSWILDQFPKWFPIERETYLDRLSLRYEREGE 310

Qy        346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405

Db        311 PNMLAPVDVFVSTVDPLKEPPLVTSNTVLSILAMDYPVEKISCYVSDDGASMLTFESLSE 370

Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465

Db        371 TAEFARKWVPFCKKFSIEPRAPEMYFTLKVDYLQDKVHPTFVKERRAMKREYEEFKVRIN 430

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525

Db        431 AQVAKASKVPLEGWIMQDGTPWPGNNTKDHPGMIQVFLGHSGGFDVEGHELPRLVYVSRE 490

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585

Db        491 KRPGFQHHKKAGAMNALVRVAGVLTNAPFMLNLDCDHYVNNSKAVREAMCFLMDPQIGKK 550

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645

Db        551 VCYVQFPQRFDGIDTNDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFKRQALYGYEPP 610

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705

Db        611 KGPKRPKMIS------CGCCPCFGRRRKNKK-----------FSKNDMNGDVAALG--- 649

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765

Db        650 ----GAEGDKEHLMFEMNFEKTFGQSSIFVTSTLMEEGGVPPSSSPAVLLKEAIHVISCG 705

Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825

Db        706 YEDKTEWGTELGWIYGSITEDILTGFKMHCRGWRSIYCMPKRPAFKGSAPINLSDRLNQV 765

Qy        826 LRWALGSIEIFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLL 884

Db        766 LRWALGSVEIFFSRHSPLWYGYKGGKLKWLERFAYANTTIYPFTSIPLLAYCILPAICLL 825

Qy        885 TGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFA 944

Db        826 TDKFIMPPISTFASLFFISLFMSIIVTGILELRWSGVSIEEWWRNEQFWVIGGISAHLFA 885

Qy        945 VFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004

Db        886 VVQGLLKILAGIDTNFTVTSKATDDDDFGELYAFKWTTLLIPPTTVLIINIVGVVAGISD 945

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064

Db        946 AINNGYQSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPTIVVIWSVLLASIFSLLWV 1005

Qy       1065 RIDPFLAKDDGPLLEECGLDC 1085

Db       1006 RIDPFVLKTKGPDTSKCGINC 1026
```

RESULT 14
Q8GSW2
ID Q8GSW2 PRELIMINARY; PRT; 1032 AA.
AC Q8GSW2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cellulose synthase.
GN CESA2.
OS Populus tremuloides (Quaking aspen).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Malpighiales; Salicaceae; Populus.
OX   NCBI_TaxID=3693;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Xylem;
RX   MEDLINE=22271539; PubMed=12383501;
RA   Samuga A., Joshi C.P.;
RT   "A new cellulose synthase gene (PtrCesA2) from aspen xylem is
RT   orthologous to Arabidopsis AtCesA7 (irx3) gene associated with
RT   secondary cell wall synthesis.";
RL   Gene 296:37-44(2002).
DR   EMBL; AY095297; AAM26299.1; -.
SQ   SEQUENCE   1032 AA;  116958 MW;  15ABB28FDC60B96D CRC64;

  Query Match             66.8%;  Score 3955.5;  DB 10;  Length 1032;
  Best Local Similarity   66.4%;  Pred. No. 0;
  Matches  729;  Conservative  138;  Mismatches  152;  Indels   79;  Gaps   15;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |||||||||||||||||||    |    || :  :||||:||||::||   || ||||||
Db          1 MEASAGLVAGSHNRNELVVIH--GHEEHKPLKNLDGQVCEICGDEIGLTVDGDLFVACNE 58

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              | ||||| |||||||||||||||||||||||||  || ||:||| |||:::||  :
Db         59 CGFPVCRPCYEYERREGTQNCPQCKTRYKRLKGSPRVEGDDEEDDVDDIEHEFIIEDEQD 118

Qy        121 QS--VAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMG 178
              ::  : |:||:| |:|||| |        : |   | :  |     :  |         :|
Db        119 KNKHLTEAMLHGKMTYGRGHDD-------EENSQFPPVITGIRSRPVSGE------FSIG 165

Qy        179 GGG---------KRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229
                |         ||:|| | ::|          |  |:          |||||: ||
Db        166 SHGEQMLSSSLHKRVHPYPVSEPG----SARWDEKKE--------GGWKERMDEWKM--- 210

Qy        230 RMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHY 289
              | || |   | |||:  ::::||| ||||:|: ||:||||||:|: ||::|   |  |
Db        211 ---QHGNLGPEQD-DDAEAAMLEDARQPLSRKVPIASSKINPYRMVIVARLIILAVFLRY 266

Qy        290 RVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQL 349
              |::|||:||  ||| |::|||||:|||||||||| ||:||||||||||:::|| |: |
Db        267 RILHPVHDALGLWLTSIVCEIWFAISWILDQFPKWLPIDRETYLDRLSLRYEQEGGPNML 326

Qy        350 APIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEF 409
              ||:| |||||||:|||||| ||:||||::||||:|:|||:|||||:| ||||:|||:||
Db        327 APVDVFVSTVDPMKEPPLVTGNTLLSILAMDYPVEKISCYLSDDGASMCTFEAMSETAEF 386

Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469
              |:|||||::|||||||:||  |:|||||   ||:||||||||||||||||:||
Db        387 ARKWVPFCKKFNIEPRAPEFYFTLKVDYLKDKVQPTFVKERRAMKREYEEFKVRINAIVA 446

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529
              |||||| ||| |||||||||| |||||||||| ||| | ||||||||||||||||
Db        447 KAQKVPTEGWIMQDGTPWPGNNTRDHPGMIQVFLGHSGGHDVEGNELPRLVYVSREKRPG 506

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589
              ::|||||||||||:|| |:|||||::||||||||:|||||::|||||:||| :||:||||
Db        507 FSHHKKAGAMNALIRVLAILTNAPFMLNLDCDHYVNNSKAVREAMCFLMDPQIGKRVCYV 566

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPK-TK 648
              |||||||||| |||||||| |||||||||||||||:||||||:|||||||  ||   |
Db        567 QFPQRFDGIDTHDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFKRQALYGYDPPKDPK 626

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
              :|   ||:       || ||| ||:|     ||                    | : |
Db        627 RPKMETCD-------CCPCFGRRKKKNA---KT-------------------GAVVEGM 656

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
                |:|  :::    |||||||::||||||:| ||   |:|||:|||||||||||||||
Db        657 D--NNDKELLMSHMNFEKKFGQSAIFVTSTLMEEGGVPPSSSPAALLKEAIHVISCGYED 714

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
              ||:|| |:|||||||||||||||| |||||||:||| |||||||:||||||:|||||
Db        715 KTEWGLELGWIYGSITEDILTGFKMHCRGWRSIYCMPKRAAFKGSAPINLSDRLNQVLRW 774

Qy        829 ALGSIEIFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGK 887
              ||||:||||| | | ||||  | ||:||||:|:|: :||:||: |:||| ||||||||  |
Db        775 ALGSVEIFFSGHSPNWYGYKKGKLKWLERFAYVNTTIYPFTSLALVAYCCLPAICLLTDK 834

Qy        888 FITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQ 947
              || ||::  |||:|::||: ||:| |||:|||||  |::|||||||||||||:|||||  |
Db        835 FIMPEISTFASLFFIALFLSIFSTGILELRWSGVSIEEWWRNEQFWVIGGVSAHLFAVVQ 894

Qy        948 GLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAIN 1007
              ||||:||:| :|||||  ||::| ||| |||||||||||:|::| :|||||||:|||
Db        895 GLLKVLAGIDLNFTVTSKATDDDDFGELYAFKWTTLLIPPTTILIINLVGVVAGVSDAIN 954

Qy       1008 NGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRID 1067
              |||:||||||||||||||||||||||||:||||||||::||:|||||||||||||
Db        955 NGYQSWGPLFGKLFFAFWVIVHLYPFLKGLMGRQNRTPTIVVIWSVLLASIFSLLWVRID 1014

Qy       1068 PFLAKDDGPLLEECGLDC 1085
              ||: |  ||  ::|||:|
Db       1015 PFVMKTRGPDTKQCGLNC 1032
```

```
RESULT 15
O48946
ID   O48946      PRELIMINARY;      PRT;  1081 AA.
AC   O48946;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Cellulose synthase catalytic subunit.
GN   RSW1 OR F8B4.110 OR AT4G32410.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98111412; PubMed=9445479;
RA   Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W.,
RA   Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
RA   Redmond J., Williamson R.E.;
RT   "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
RL   Science 279:717-720(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,
RA   De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA   Villaroel R., Gielen J., Van Montagu M., Hoheisel J., Mewes H.W.,
RA   Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Terryn N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,
RA   De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,
RA   Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF027172; AAC39334.1; -.
DR   EMBL; AL034567; CAA22568.1; -.
DR   EMBL; AL161581; CAB79958.1; -.
DR   InterPro; IPR005150; Cellulose_synt.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF03552; Cellulose_synt; 1.
```

```
DR   SMART; SM00184; RING; 1.
SQ   SEQUENCE   1081 AA;  122236 MW;  BDEB5D9DEE334D59 CRC64;

  Query Match             66.7%;  Score 3945.5;  DB 10;  Length 1081;
  Best Local Similarity   66.9%;  Pred. No. 0;
  Matches  727;  Conservative  140;  Mismatches  182;  Indels   37;  Gaps   11;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |||||||||||: ||||| || : | | || :  |||:|||||||||||  || ||||||
Db          1 MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMNGQICQICGDDVGLAETGDVFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||| ||||||::||| |||||||::| :|  || |||:|| |||::||||:    :
Db         61 CAFPVCRPCYEYERKDGTQCCPQCKTRFRRHRGSPRVEGDEDEDDVDDIENEFNY----A 116

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDD---IPPEQHALVPSF- 176
              |   ::       : | |:   :    : .| :||||:|  |       |  |      |
Db        117 QGANKA------RHQRHGEEFSSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER--MHQT 234
              :|   :      || ||  ||   | :|||||| :|| |:| ||||:| || :||:  :  |
Db        170 LGPSDRNAISSPYIDPRQPVPVRIVDPSKDLNSYGLGNVDWKERVEGWKLKQEKNMLQMT 229

Qy        235 GN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285
              |          :| | :|:  :| : |: |  :|| :|:|||:: |||::||:||::| |
Db        230 GKYHEGKGGEIEGTGSNGE--ELQMADDTRLPMSRVVPIPSSRLTPYRVVIILRLIILCF 287

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQ 345
              |  ||  ||| :|: ||| ||||||||| ||:|||||||:|| |||||||::|:|::|:
Db        288 FLQYRTTHPVKNAYPLWLTSVICEIWFAFSWLLDQFPKWYPINRETYLDRLAIRYDRDGE 347

Qy        346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405
              |||| |:| |||||||||||||| |||||||||||||||:|||||||:||||||:|||
Db        348 PSQLVPVDVFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVACYVSDDGSAMLTFESLSE 407

Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465
              |:||||||||||::|||||||:|| |||||||||:  :||:||||||||||||||||||
Db        408 TAEFAKKWVPFCKKFNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKREYEEFKVRIN 467

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525
              ||||||||:|||||||||||||||| |||||||||||| |||||| :||||||:|||||
Db        468 ALVAKAQKIPEEGWTMQDGTPWPGNNTRDHPGMIQVFLGHSGGLDTDGNELPRLIYVSRE 527

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585
              ||||: |||||||||||:||||||||  ||||:||||| |||||||||||||||| :|||
Db        528 KRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK 587

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645
              ||||||||||||| ||||||||:|||||||||||||||:|||||| | ||||||||
Db        588 CCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFNRQALYGYDPV 647

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705
              |::      :  |       ||   :| | : |   ||:: :   : :: :| : : :||
Db        648 LTEE------DLEPNIIVKSCCGSRKKGKSSKKYNYEKRRGI--NRSDSNAPLFNMEDID 699

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765
              |    | ::|:: :::|: :||:|||| ||: :| :| ||    : :||:||||||||||||
Db        700 EGFEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIPPTTNPATLLKEAIHVISCG 759

Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825
              ||||||||||||||||||||||||||  || |||| | | |||||||:||||||:||
Db        760 YEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCNPPRPAFKGSAPINLSDRLNQV 819

Qy        826 LRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLT 885
              ||||||||||   | |||:|||| | |: ||| :|||:|||| |||||:||| ||| ||:|
Db        820 LRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINTIVYPITSIPLIAYCILPAFCLIT 879

Qy        886 GKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAV 945
              :|| ||::| ||:||: ||| |  | |||:||||| |:|||||||||||| |:|||||
Db        880 DRFIIPEISNYASIWFILLFISIAVTGILELRWSGVSIEDWWRNEQFWVIGGTSAHLFAV 939

Qy        946 FQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004
              |||||||:||:||:||||||  | | :|:||| |||| |||||||:||:| ||:|||||
Db        940 FQGLLKVLAGIDTNFTVTSKATDEDGDFAELYIFKWTALLIPPTTVLLVNLIGIVAGVSY 999

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
              |:|:||:|||||||||||| ||| |||||||||:|||||||||||||:||||||||||
Db       1000 AVNSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPTIVIVWSVLLASIFSLLWV 1059

Qy       1065 RIDPFL 1070
              ||:||:
Db       1060 RINPFV 1065


Search completed: December 15, 2003, 13:04:19
Job time : 51 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:56:44 ; Search time 46 Seconds
(without alignments)
3747.327 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5918	100.0	1086	21	AAY84111	Amino acid sequenc
2	5918	100.0	1086	21	AAY84116	Amino acid sequenc
3	5918	100.0	1086	21	AAY58835	Corn cellulose syn
4	5570.5	94.1	1148	21	AAY58832	Corn cellulose syn
5	5402.5	91.3	1058	21	AAY84107	Amino acid sequenc
6	5228	88.3	1094	21	AAY84113	Amino acid sequenc
7	5228	88.3	1094	21	AAY84118	Amino acid sequenc
8	5228	88.3	1165	21	AAY58834	Corn cellulose syn
9	4278.5	72.3	1084	23	ABB93304	Herbicidally activ
10	4262	72.0	1084	23	ABB93949	Herbicidally activ
11	4250.5	71.8	1084	19	AAW33818	Arabidopsis cellul
12	4209.5	71.1	1088	23	ABB91831	Herbicidally activ
13	4204.5	71.0	1069	23	ABB93411	Herbicidally activ
14	4185.5	70.7	1080	22	AAB37895	Arabidopsis thalia
15	3986	67.4	1039	21	AAY58837	Soybean cellulose
16	3984	67.3	1075	21	AAY84109	Amino acid sequenc
17	3984	67.3	1075	21	AAY84115	AMino acid sequenc
18	3984	67.3	1075	21	AAY84120	Amino acid sequenc
19	3977.5	67.2	1074	21	AAY84110	Amino acid sequenc
20	3977.5	67.2	1074	21	AAY84112	Amino acid sequenc
21	3977.5	67.2	1074	21	AAY84117	Amino acid sequenc
22	3967.5	67.0	1026	23	ABB93522	Herbicidally activ
23	3945.5	66.7	1081	19	AAW33817	Arabidopsis cellul
24	3945.5	66.7	1081	23	ABB93199	Herbicidally activ
25	3935.5	66.5	1081	19	AAW33820	Arabidopsis cellul
26	3911	66.1	1076	21	AAY84121	DNA encoding a mai
27	3903.5	66.0	1039	19	AAW73309	Cellulose synthase
28	3899.5	65.9	1077	21	AAY84108	Amino acid sequenc
29	3899.5	65.9	1077	21	AAY84114	Amino acid sequenc
30	3898.5	65.9	1065	21	AAG48733	Arabidopsis thalia
31	3898.5	65.9	1065	23	ABB93356	Herbicidally activ
32	3887.5	65.7	1054	21	AAG48734	Arabidopsis thalia
33	3880.5	65.6	1079	23	AAU10496	Corn cellulose syn
34	3876.5	65.5	1065	19	AAW33819	Arabidopsis cellul
35	3804.5	64.3	1043	23	ABB93701	Herbicidally activ
36	3782	63.9	1065	23	ABB91880	Herbicidally activ
37	3780.5	63.9	1043	21	AAY84119	Amino acid sequenc
38	3603	60.9	959	21	AAG48735	Arabidopsis thalia
39	3521	59.5	793	21	AAY58839	Soybean cellulose
40	3476.5	58.7	974	19	AAW73308	Cellulose synthase
41	3469.5	58.6	974	19	AAW60039	Cotton cellulose s
42	3356	56.7	955	21	AAB16338	Eucalyptus grandis
43	3329.5	56.3	958	23	ABB92981	Herbicidally activ
44	3318	56.1	821	23	AAU10495	Corn cellulose syn
45	3138.5	53.0	946	22	AAB37893	Populus tremuloide

ALIGNMENTS

RESULT 1
AAY84111
ID AAY84111 standard; Protein; 1086 AA.
XX
AC AAY84111;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR N-PSDB; AAZ99503.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants

PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 15; Page 118-120; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 1086 AA;

Query Match 100.0%; Score 5918; DB 21; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Db    1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy   61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
Db   61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy  121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
Db  121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy  181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
Db  181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy  241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
Db  241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy  301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
Db  301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy  361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
Db  361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy  421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
Db  421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy  481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
Db  481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy  541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
Db  541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy  601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
Db  601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy  661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
Db  661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy  721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
Db  721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy  781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840
Db  781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy  841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900
Db  841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy  901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960
Db  901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy  961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020
Db  961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy 1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080
Db 1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy 1081 CGLDCN 1086
Db 1081 CGLDCN 1086
```

RESULT 2
AAY84116
ID AAY84116 standard; Protein; 1086 AA.
XX
AC AAY84116;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR N-PSDB; AAZ99518.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 15; Page 157-160; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.

```
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence   1086 AA;

  Query Match             100.0%;  Score 5918;  DB 21;  Length 1086;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1086;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy       1081 CGLDCN 1086
              ||||||
Db       1081 CGLDCN 1086


RESULT 3
AAY58835
ID   AAY58835 standard; Protein; 1086 AA.
XX
AC   AAY58835;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Corn cellulose synthase (full-length).
XX
KW   Corn; maize; cellulose synthase; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO200004166-A2.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15871.
XX
PR   14-JUL-1998;   98US-0092844.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Fader GM,  Falco SC,  Kinney AJ,  Lightner JE,  Miao G;
PI   Rafalski JA,  Thorpe CJ;
XX
DR   WPI; 2000-182431/16.
DR   N-PSDB; AAZ58266.
XX
PT   Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT   probes for isolating cDNAs and genes encoding homologous proteins, for
PT   producing transgenic plants  -
XX
PS   Claim 12; Page 50-53; 93pp; English.
XX
CC   The present sequence is that of full-length corn cellulose
CC   synthase (CS), as deduced from an isolated cDNA clone (see AAY58266).
CC   The invention relates to isolated CS nucleic acid fragments and to
CC   CS polypeptides.  It also relates to the construction of a chimeric
CC   gene encoding all or a portion of the CS, in sense or antisense
CC   orientation, where expression of the gene results in altered levels
```

```
CC   of the CS in transformed host cells.  The host cells can be used to
CC   screen compounds for their ability to inhibit CS activity.  CS
CC   nucleic acids are also useful for producing transgenic plants having
CC   altered levels of CS, and hence altered levels of fibre.  CS may
CC   also serve as a target for the development of novel herbicides.
XX
SQ   Sequence   1086 AA;

  Query Match             100.0%;  Score 5918;  DB 21;  Length 1086;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1086;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy       1081 CGLDCN 1086
              ||||||
Db       1081 CGLDCN 1086


RESULT 4
AAY58832
ID   AAY58832 standard; Protein; 1148 AA.
XX
AC   AAY58832;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Corn cellulose synthase (partial sequence).
XX
KW   Corn; maize; cellulose synthase; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO200004166-A2.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15871.
XX
PR   14-JUL-1998;   98US-0092844.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Fader GM,  Falco SC,  Kinney AJ,  Lightner JE,  Miao G;
PI   Rafalski JA,  Thorpe CJ;
XX
DR   WPI; 2000-182431/16.
DR   N-PSDB; AAZ58263.
XX
PT   Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT   probes for isolating cDNAs and genes encoding homologous proteins, for
PT   producing transgenic plants  -
XX
PS   Claim 12; Page 39-42; 93pp; English.
XX
CC   The present sequence is that of a portion of corn cellulose
CC   synthase (CS), as deduced from a contig of cDNA clones (see AAY58263).
CC   The invention relates to isolated CS nucleic acid fragments and to
CC   CS polypeptides.  It also relates to the construction of a chimeric
CC   gene encoding all or a portion of the CS, in sense or antisense
CC   orientation, where expression of the gene results in altered levels
CC   of the CS in transformed host cells.  The host cells can be used to
CC   screen compounds for their ability to inhibit CS activity.  CS
CC   nucleic acids are also useful for producing transgenic plants having
CC   altered levels of CS, and hence altered levels of fibre.  CS may
CC   also serve as a target for the development of novel herbicides.
XX
SQ   Sequence   1148 AA;

  Query Match             94.1%;  Score 5570.5;  DB 21;  Length 1148;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches 1027;  Conservative   23;  Mismatches   35;  Indels    5;  Gaps    3;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |||||||||||||||||||||||:|||||  ::|||||||||||||  | |:|||||||
Db         60 MEASAGLVAGSHNRNELVVIRRDGEPGPKPMDQRNGQVCQICGDDVGRNPDGEPFVACNE 119
```

```
Qy        61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW-DGHD 119
             ||||:|||||||||||||||||||||:|||||| || ||||||||||:|||||  |  ||
Db       120 CAFPICRDCYEYERREGTQNCPQCKTRFKRLKGCARVPGDEEEDGVDDLENEFNWSDKHD 179

Qy       120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
             || :|||||: |||||||  | :|  || |   ||||||||||||||||:|||||||:||
Db       180 SQYLAESMLHAHMSYGRGADLDGVPQPFHPIPNVPLLTNGQMVDDIPPDQHALVPSFVGG 239

Qy       180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGG 239
             |||||||||||||:|||||||||||||||||||||||||||||:|||:|||||| ||||
Db       240 GGKRIHPLPYADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMESWKQKQERMHQTRNDGG 299

Qy       240 GDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAF 299
             |||||||||||||||| |||||||||||||||||||||||||:|||:|||||||  ||||
Db       300 GDDGDDADLPLMDEARQPLSRKIPLPSSQINPYRMIIIIRLVVLCFFFHYRVMHPVPDAF 359

Qy       300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTV 359
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||:|||||||
Db       360 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGHPSQLAPVDFFVSTV 419

Qy       360 DPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 419
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       420 DPLKEPPLVTANTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 479

Qy       420 YNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 479
             |::|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       480 YSLEPRAPEWYFQQKIDYLKDKVAPNFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 539

Qy       480 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAM 539
             |||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db       540 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDVEGNELPRLVYVSREKRPGYNHHKKAGAM 599

Qy       540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       600 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 659

Qy       600 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       660 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 719

Qy       660 KWCFCCCCFGNRKQK---KTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKA 716
             ||| |||||||| |    ||:||| || | | |||:|||:|||||||||||||||||||
Db       720 KWCICCCCFGNRKTKKKTKTSKPKFEKIKKL-FKKKENQAPAYALGEIDEAAPGAENEKA 778

Qy       717 GIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEI 776
              |||||||||||||||| |||||||||||||||||||||||||||||||| |||:|
Db       779 SIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYEDKTGWGKDI 838

Qy       777 GWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIF 836
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       839 GWIYGSVTEDILTGFKMHCHGWRSIYCIPKRAAFKGSAPLNLSDRLHQVLRWALGSIEIF 898

Qy       837 FSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNV 896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       899 FSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNV 958

Qy       897 ASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGV 956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       959 ASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGV 1018

Qy       957 DTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPL 1016
             ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db      1019 DTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINNGYESWGPL 1078

Qy      1017 FGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGP 1076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1079 FGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGP 1138

Qy      1077 LLEECGLDCN 1086
             ||||||||||
Db      1139 LLEECGLDCN 1148
```

RESULT 5
AAY84107
ID AAY84107 standard; Protein; 1058 AA.
XX
AC AAY84107;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR N-PSDB; AAZ99491.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 15; Page 87-89; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 1058 AA;

Query Match 91.3%; Score 5402.5; DB 21; Length 1058;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 994; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

```
Qy        33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLK 92
             ::|||||||||||||  | |:||||||||||:|||||||||||||||||||:|| |
Db         3 QRNGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYEYERREGTQNCPQCKTRFKRFK 62

Qy        93 GCQRVTGDEEEDGVDDLDNEFNW-DGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNP 151
             || || |||||||||||:||||| | |||| :|||||: |||||||  | :|  || |   |
Db        63 GCARVPGDEEEDGVDDLENEFNWSDKHDSQYLAESMLHAHMSYGRGADLDGVPQPFHPIP 122
```

```
Qy        152 NVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYG 211
              ||||||||||||||||:||||||||:||||||||||||||||:||||||||||||||||||
Db        123 NVPLLTNGQMVDDIPPDQHALVPSFVGGGGKRIHPLPYADPNLPVQPRSMDPSKDLAAYG 182

Qy        212 YGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINP 271
              |||||||||||:|||:|||||||| |||||||||||||||||||| ||||||||||||||
Db        183 YGSVAWKERMESWKQKQERMHQTRNDGGGDDGDDADLPLMDEARQPLSRKIPLPSSQINP 242

Qy        272 YRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 331
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||
Db        243 YRMIIIIRLVVLCFFFHYRVMHPVPDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 302

Qy        332 YLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVS 391
              |||||||||||| ||||||:|||||||||||||||| ||||||||||||||||||||||||
Db        303 YLDRLSLRFDKEGHPSQLAPVDFFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVSCYVS 362

Qy        392 DDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERR 451
              ||||||||||||||||||||||||||::|||||||||||||||||||| |||||||||||
Db        363 DDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQQKIDYLKDKVAPNFVRERR 422

Qy        452 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDC 511
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        423 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDV 482

Qy        512 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 542

Qy        572 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 602

Qy        632 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQK---KTTKPKTEKKKLLF 688
              |||||||||||||||||||||||||||||| ||||||||| |   ||:||| || | |
Db        603 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCICCCCFGNRKTKKKTKTSKPKFEKIKKL- 661

Qy        689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748
              |||:|||:|||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        662 FKKKENQAPAYALGEIDEAAPGAENEKASIVNQQKLEKKFGQSSVFVASTLLENGGTLKS 721

Qy        749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
              |||||||||||||||||||| |||:|||||||||||||||||||||||||||||||||||
Db        722 ASPASLLKEAIHVISCGYEDKTGWGKDIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRA 781

Qy        809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
              |||||||||||| |||||||||||| |||||||||||||||||||||||||||||||||||
Db        782 AFKGSAPLNLSDRFHQVLRWALGSIEILFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 841

Qy        869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 901

Qy        929 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 961

Qy        989 TLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1048
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TLLLLNFIGVVAGISNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1021

Qy       1049 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDC 1085
              |||||||||||||||||||||||||||||||||||||
Db       1022 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDC 1058
```

RESULT 6
AAY84113
ID AAY84113 standard; Protein; 1094 AA.
XX
AC AAY84113;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR N-PSDB; AAZ99509.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 15; Page 134-136; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 1094 AA;

Query Match 88.3%; Score 5228; DB 21; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57
              |||||||||||||||||||||||   |  |   |     ||||||:||:   |:||||
Db          1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116
              |||||||||| |||||||||:| ||||:|||||||| ||.||||||||||: ||   |
Db         60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172
              |    | | ||||||   ||||||||   |   |   ||||||||||||||||||||||
Db        120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229
              |||:|   ||||||||||:|||:||||||||||||||||||||||||||||| |||:||
Db        177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFH 288
```

```
Db        237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348

Db        297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408

Db        357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468

Db        417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528

Db        477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588

Db        537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596

Qy        589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648

Db        597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708

Db        657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768

Db        717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828

Db        777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888

Db        837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948

Db        897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008

Db        957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068

Db       1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy       1069 FLAKDDGPLLEECGLDCN 1086

Db       1077 FLAKSNGPLLEECGLDCN 1094


RESULT 7
AAY84118
ID   AAY84118 standard; Protein; 1094 AA.
XX
AC   AAY84118;
XX
DT   03-JUL-2000  (first entry)
XX
DE   Amino acid sequence of a maize cellulose synthase.
XX
KW   Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW   transgenic plant; plant breeding marker.
XX
OS   Zea mays.
XX
PN   WO200009706-A2.
XX
PD   24-FEB-2000.
XX
PF   16-AUG-1999;   99WO-US18760.
XX
PR   17-AUG-1998;   98US-0096822.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Dhugga KS,  Helentjaris TG,  Bowen BA,  Wang X;
XX
DR   WPI; 2000-224343/19.
DR   N-PSDB; AAZ99524.
XX
PT   New genes which encode maize cellulose synthase polypeptides in plants
PT   useful for modulating the expression of cellulose synthase in plants
PT   and to produce transgenic plants expressing the novel protein  -
XX
PS   Claim 15; Page 173-175; 119pp; English.
XX
CC   The present sequence represents a maize cellulose synthase polypeptide.
CC   The cellulose synthase can be used for the improvement of stalk quality
CC   for improved stand or silage. It also provides an increased concentration
CC   of cellulose in the pericarp, hardening the kernel and improving its
CC   handling ability. The sequences are used to produce transgenic plants
CC   and seeds expressing the cellulose synthase. The polynucleotide is
CC   used for modulating, preferably increasing, the level of the synthase
CC   in a plant cell. The plants are preferably monocots. The polynucleotide
CC   is also used as a probe or primer in the detection quantitation or
CC   isolation of gene transcripts. The probes are useful in detecting
CC   deficiencies in the level of mRNA in screenings for desired transgenic
CC   plant, for detecting mutations in the gene, for monitoring upregulation
CC   of expression or changes in enzyme activity in screening assays of
CC   compounds, for detection of any number of allelic variants of the gene,
CC   or for use as molecular markers in plant breeding programs. The
CC   isolated nucleic acids of the present invention can also be used for
CC   recombinant expression of their encoded polypeptides or for use as
CC   immunogens in the preparation and/or screening of antibodies. The
CC   proteins can be employed in assays for enzyme agonists or antagonists
CC   of enzyme function or for use of immunogens or antigens to obtain
CC   antibodies specifically immunoreactive with a protein.
XX
SQ   Sequence   1094 AA;

  Query Match             88.3%;  Score 5228;  DB 21;  Length 1094;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57

Db          1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116

Db         60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172

Db        120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229

Db        177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFH 288

Db        237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348
```

```
Db        297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408

Db        357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468

Db        417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528

Db        477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588

Db        537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596

Qy        589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648

Db        597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708

Db        657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768

Db        717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828

Db        777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888

Db        837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948

Db        897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008

Db        957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068

Db       1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy       1069 FLAKDDGPLLEECGLDCN 1086

Db       1077 FLAKSNGPLLEECGLDCN 1094


RESULT 8
AAY58834
ID   AAY58834 standard; Protein; 1165 AA.
XX
AC   AAY58834;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Corn cellulose synthase (partial sequence).
XX
KW   Corn; maize; cellulose synthase; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO200004166-A2.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15871.
XX
PR   14-JUL-1998;   98US-0092844.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Fader GM,  Falco SC,  Kinney AJ,  Lightner JE,  Miao G;
PI   Rafalski JA,  Thorpe CJ;
XX
DR   WPI; 2000-182431/16.
DR   N-PSDB; AAZ58265.
XX
PT   Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT   probes for isolating cDNAs and genes encoding homologous proteins, for
PT   producing transgenic plants  -
XX
PS   Claim 12; Page 45-49; 93pp; English.
XX
CC   The present sequence is that of a portion of corn cellulose
CC   synthase (CS), as deduced from an isolated cDNA clone (see AAY58265).
CC   The invention relates to isolated CS nucleic acid fragments and to
CC   CS polypeptides.  It also relates to the construction of a chimeric
CC   gene encoding all or a portion of the CS, in sense or antisense
CC   orientation, where expression of the gene results in altered levels
CC   of the CS in transformed host cells.  The host cells can be used to
CC   screen compounds for their ability to inhibit CS activity.  CS
CC   nucleic acids are also useful for producing transgenic plants having
CC   altered levels of CS, and hence altered levels of fibre.  CS may
CC   also serve as a target for the development of novel herbicides.
XX
SQ   Sequence   1165 AA;

  Query Match           88.3%;  Score 5228;  DB 21;  Length 1165;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57

Db         72 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 130

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116

Db        131 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 190

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172

Db        191 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 247

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229

Db        248 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 307

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFH 288

Db        308 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 367

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348

Db        368 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 427

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408

Db        428 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 487

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468

Db        488 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 547

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528

Db        548 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 607

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588
```

```
                 |||||||||||||||||||||:|| |||||||||||||||||||||||||||||:||||||
Db        608 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 667

Qy        589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648
              ||||||||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 727

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
              |||||||||||||  |||  |: :||||||||||||| ||||| || |||||||||||| |
Db        728 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 787

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
              |||: |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        788 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 847

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
              |||||||||||||:|||||||||||||||||||| ||||||||||||||||||||||||
Db        848 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 907

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
              ||||:||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 967

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
              ||||| ||||:|||:|||||  | |||||||| ||||||||||||||:||||||||
Db        968 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 1027

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
              |||| ||:|||||||||| ||||||||||||||||||||||||||||||:||||||
Db       1028 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1087

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       1088 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1147

Qy       1069 FLAKDDGPLLEECGLDCN 1086
              |||| :||||||||||||
Db       1148 FLAKSNGPLLEECGLDCN 1165


RESULT 9
ABB93304
ID   ABB93304 standard; Protein; 1084 AA.
XX
AC   ABB93304;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 2515.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 2515; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   1084 AA;

  Query Match             72.3%;  Score 4278.5;  DB 23;  Length 1084;
  Best Local Similarity   71.3%;  Pred. No. 0;
  Matches  782;  Conservative  127;  Mismatches  161;  Indels   27;  Gaps   12;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:|||||||| |:|  |      :  :| :|| ||||||:: |    : ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
              ||||||| ||||||||| | ||||||||||:||  || | ||||: :|||: ||:  | |
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEFD-HGMD 119

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
               :  ||:|   :: ||||  :  |         :||||       |:  ::|||:     |
Db        120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
              | |::| |: | | | | ||| | ||:| |||||||||:||| ||:|| |::    |:
Db        175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy        235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290
              ||:| |     |: || |:|:||| || ||||:|: ||:||||||:|: || :|| |||||
Db        235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy        291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350
              ::||||||: ||| |||||||||:||||||||||:|||||||||||||::|||:|| ||
Db        295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy        351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410
              |:| |||||||||||||:| |||||||:|||||||:|||||||||||||||||:|:|||
Db        355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSDTAEFA 414

Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470
              :|||||||::|||||||||| ||:||||:||   ||||||||||:|||||||
Db        415 RKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530
              |||||||||||||||||||||||||||||||| ||   | :||||||||||||||||:
Db        475 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590
              :|||||||:|:||||||:|||||||:||||||||||:|:|||||||  ||||||||
Db        535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
              |||||||||||||:|||||||||||||||||||||||||||||||||||||:|||| |||
Db        595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKKP 654

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEA--A 708
              | :||||||||| | || | ||: ||    | : ||      :|     :||  :||
Db        655 PGKTCNCWPKWC-CLCC-GLRKKSKT---KAKDKKT---NTKETSKQIHALENVDEGVIV 706

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
              | :  ||      | |||||||| ||| | :|:|||  ::|||| ||:||| ||||||||
Db        707 PVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGYED 766

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
              ||:||||||||||||||||||||||||:||:||| |||||||:||||||||||||||
Db        767 KTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRW 826
```

```
Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
Db        827 ALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTGKF 886

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
Db        887 IVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALFQG 946

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
Db        947 LLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAISN 1006

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
Db       1007 GYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRVNP 1066

Qy       1069 FLAKDDGPLLEECGLDC 1085
Db       1067 FVAK-GGPVLEICGLNC 1082
```

RESULT 10
ABB93949
ID ABB93949 standard; Protein; 1084 AA.
XX
AC ABB93949;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 3160.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 3160; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 1084 AA;

Query Match 72.0%; Score 4262; DB 23; Length 1084;
Best Local Similarity 71.1%; Pred. No. 0;
Matches 778; Conservative 126; Mismatches 170; Indels 20; Gaps 10;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Db          1 MNTGGRLIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW--DGH 118
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDDIDDLDNEFEYGNNGI 120

Qy        119 DSQSVAESMLYGHMSYG-RGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL-VPSF 176
Db        121 GFDQVSEGMSISRRNSGFPQSDLDSAPPGSQ----IPLLTYGDEDVEISSDRHALIVPPS 176

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERMHQTG 235
Db        177 LGGHGNRVHPVSLSDPTVAAHPRPMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVR 236

Qy        236 NDGGGD--DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMH 293
Db        237 HEGDPDFEDGDDADFPMMDEGRQPLSRKIPIKSSKINPYRMLIVLRLVILGLFFHYRILH 296

Qy        294 PVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPID 353
Db        297 PVKDAYALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVD 356

Qy        354 FFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKW 413
Db        357 VFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSETAEFARKW 416

Qy        414 VPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQK 473
Db        417 VPFCKKYCIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQK 476

Qy        474 VPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHH 533
Db        477 VPEDGWTMQDGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPGFDHH 536

Qy        534 KKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQ 593
Db        537 KKAGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQ 596

Qy        594 RFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSR 653
Db        597 RFDGIDRHDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKGPRK 656

Qy        654 TCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDE--AAPGA 711
Db        657 TCNCWPKWCL--LCFGSRKNRKAKTVAADKKK----KNREASKQIHALENIEEGRVTKGS 710

Qy        712 ENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTD 771
Db        711 NVEQSTEAMQMKLEKKFGQSPVFVASARMENGGMARNASPACLLKEAIQVISCGYEDKTE 770

Qy        772 WGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALG 831
Db        771 WGKEIGWIYGSVTEDILTGFKMHSHGWRSVYCTPKLAAFKGSAPINLSDRLHQVLRWALG 830

Qy        832 SIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITP 891
Db        831 SVEIFLSRHCPIWYGYGGGLKWLERLSYINSVVYPWTSLPLIVYCSLPAICLLTGKFIVP 890

Qy        892 ELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLK 951
Db        891 EISNYASILFMALFSSIAITGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLK 950

Qy        952 VIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYE 1011
Db        951 VLAGVDTNFTVTSKAADDGEFSDLYLFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYD 1010

Qy       1012 SWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLA 1071
Db       1011 SWGPLFGRLFFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVA 1070

Qy       1072 KDDGPLLEECGLDC 1085
```

```
                |  ||:|| |||||
Db       1071 K-GGPILEICGLDC 1083


RESULT 11
AAW33818
ID   AAW33818 standard; Protein; 1084 AA.
XX
AC   AAW33818;
XX
DT   06-JUL-1998  (first entry)
XX
DE   Arabidopsis cellulose biosynthetic protein Ath-A.
XX
KW   Cellulose; cellulose synthase; RSW1 protein; beta-1,4-glucan;
KW   transgenic plant.
XX
OS   Arabidopsis thaliana var. Columbia.
XX
PN   WO9800549-A1.
XX
PD   08-JAN-1998.
XX
PF   24-JUN-1997;   97WO-AU00402.
XX
PR   27-JUN-1996;   96AU-0000699.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Arioli A,  Betzner AS,  Peng L,  Williamson RE;
XX
DR   WPI; 1998-086974/08.
DR   N-PSDB; AAV06566.
XX
PT   DNA encoding cellulose biosynthetic enzyme - useful for manipulation
PT   of cellulose and beta-1,4-glucan
XX
PS   Claim 29; Page 130-135; 207pp; English.
XX
CC   This polypeptide is encoded by cDNA clone Ath-A (see AAV06566)
CC   that was isolated from an Arabidopsis thaliana cDNA library using
CC   PCR primers (see AAT99632-34) based on cellulose synthase RSW1
CC   genomic clone 23H12 (see AAV06563) and EST clone AAT20782 (see
CC   AAV06562). It is closely related to Arabidopsis cellulose synthase RSW1
CC   protein (see AAW33816-17).  Claimed nucleic acid molecules (see
CC   AAV06562-69) coding for claimed polypeptides (see AAW33816-20 and
CC   AAW46202) involved in cellulose biosynthesis can be used to
CC   manipulate the cellulose and/or beta-glucan content of transgenic
CC   plants.  Expression of nucleic acids in the sense orientation
CC   increases the level of cellulose and reduces the level of
CC   non-crystalline beta-1,4-glucan and starch, providing plants with
CC   modified strength and/or shape and/or fibre properties, or having
CC   increased resistance to stresses or pests.  Antisense, ribozyme or
CC   co-suppression molecules can be used to reduce the cellulose
CC   content of a transgenic plant, e.g. to improve digestibility or to
CC   alter carbon partitioning such that increased carbon is available
CC   for growth, rather than deposited as cellulose.
XX
SQ   Sequence   1084 AA;

  Query Match             71.8%;  Score 4250.5;  DB 19;  Length 1084;
  Best Local Similarity   70.8%;  Pred. No. 0;
  Matches  778;  Conservative  128;  Mismatches  162;  Indels   31;  Gaps   12;

Qy           1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
               |     |:|||||||| |:|  |     :  :| :|| ||||||:: |    : ||||||
Db           1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy          61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
               ||||||| ||||||||| | ||||||||||:||  || | ||||: :|||: ||:  | |
Db          61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEFD-HGMD 119

Qy         120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
               :  ||: |   :: |||| :  |        :|||      |:  ::|||:     |
Db         120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy         180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
               | |::| |: | | | | ||| | ||:| |||||||||:||| ||:|| |::    |:
Db         175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy         235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290
               ||:| |    |: || |:|:||| || ||||:|: ||:|||||||:|: || :|| |||||
Db         235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy         291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350
               ::|||||: ||| |||||||||:|||||||||:|||||||||||||||::|||:|| ||
Db         295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy         351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410
               |:| |||||||||||||:| |||||||:||||||||:||||::|||||||||||:|::||
Db         355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSNNGAAMLTFEALSDTADFA 414

Qy         411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470
               |||||||::|||||||||| ||:||||:||   ||||||||||:|||||:||||||
Db         415 TKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy         471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530
               ||||||| |||||||||||||||||||||||||| ||   | :|||||||||||||||:
Db         475 AQKVPEERWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy         531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590
               :|||||||||:|:||||||:|||||||:|||||||||||:|:|||||  ||||||||
Db         535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy         591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650
               |||||||||||||:|||||||||||||||||||||||||||:|||||:|||| |||
Db         595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRKQALYGFDAPKKKKP 654

Qy         651 PSRTCNCWPKWCFCCCCFGNRKQKKT--TKPKTEKKKLLFFKKEENQSPAYALGEIDEA- 707
               | :||||||||| | || | ||: ||  |  ||  |      |     :||  :||
Db         655 PGKTCNCWPKWC-CLCC-GLRKKSKTKATDKKTNTK--------ETSKQIHALENVDEGV 704

Qy         708 -APGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGY 766
                | :  ||      | |||||||| ||| | :|:||| ::|||| ||:||| ||||||
Db         705 IVPVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGY 764

Qy         767 EDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVL 826
               :|||:|||||||||||||||||||||||||||:||:||| |||||||:||||||||||
Db         765 QDKTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVL 824

Qy         827 RWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTG 886
               |||||:||| | |||:|||||||:|||||||||:||||||:||: ||:|||:|||||
Db         825 RWALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTG 884

Qy         887 KFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVF 946
               ||| ||::| | : || :|| |  | ||||:| |||||||||||||||||| ||||||:|
Db         885 KFIVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALF 944

Qy         947 QGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAI 1006
               ||||||:|||:|:|||||| ||  ||||| |||||||||||||::| |||: |||:||
Db         945 QGLLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAI 1004

Qy        1007 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRI 1066
               :|||:|||||||:|||| |||||||||||::|:|:: |||::|||||||| :||||||
Db        1005 SNGYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRI 1064

Qy        1067 DPFLAKDDGPLLEECGLDC 1085
               :||:||  ||:|| |||:|
Db        1065 NPFVAK-GGPVLEICGLNC 1082


RESULT 12
ABB91831
ID   ABB91831 standard; Protein; 1088 AA.
XX
```

```
AC   ABB91831;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 1042.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 1042; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   1088 AA;

  Query Match             71.1%;  Score 4209.5;  DB 23;  Length 1088;
  Best Local Similarity   70.2%;  Pred. No. 0;
  Matches  769;  Conservative  127;  Mismatches  179;  Indels   21;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:||||||| |:|   |      :   | :|| |:|| |:: |    |:||:||||
Db          1 MNTGGRLIAGSHNRNEFVLINADDTARIRSAEELSGQTCKICRDEIELTDNGEPFIACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||  || |||||||||  | |||| |||||:||  || ||||:| :|||::||    | |
Db         61 CAFPTCRPCYEYERREGNQACPQCGTRYKRIKGSPRVEGDEEDDDIDDLEHEFY--GMDP 118

Qy        121 QSVAESML-YGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
              : | |: | |  :: ||| |      :      |||||      |:  ::|||:      |
Db        119 EHVTEAALYYMRLNTGRGTDEVSHLYSASPGSEVPLLTYCDEDSDMYSDRHALIVPPSTG 178

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
              | |:| :|: |    :  | | | |||  |||||||||:||| ||::| |::    ::
Db        179 LGNRVHHVPFTDSFASIHTRPMVPQKDLTVYGYGSVAWKDRMEVWKKQQIEKLQVVKNER 238

Qy        235 GNDGGG-----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHY 289
              ||| |     |: ||  ||:||| || ||||:|: ||:|||||:|  || :|| ||||
Db        239 VNDGDGDGFIVDELDDPGLPMMDEGRQPLSRKLPIRSSRINPYRMLIFCRLAILGLFFHY 298

Qy        290 RVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQL 349
              |::|||||| ||| |||||||||:|||||||||:|||||||||||||::|||:||:|
Db        299 RILHPVNDAFGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSEL 358

Qy        350 APIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEF 409
              ||:| |||||||||||||:| |||||||:|||||:||:|||||||||||||||||| |:||
```

```
Db        359 APVDVFVSTVDPLKEPPLITANTVLSILAVDYPVEKVACYVSDDGAAMLTFEALSYTAEF 418

Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469
              |:|||||||:::|||||||||| ||:|||| ||    || |||||||:|||||:|||||:
Db        419 ARKWVPFCKKFSIEPRAPEWYFSQKMDYLKHKVDPAFVMERRAMKRDYEEFKVKINALVS 478

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529
               :|||||:|||||||||||||||||||||||||||| ||   | :|||||||||||||||||
Db        479 VSQKVPEDGWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVCDMDGNELPRLVYVSREKRPG 538

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589
              ::|||||||||:|:|||||||:|||||||:||||||||||||:|||||||||  |||:|||
Db        539 FDHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIREAMCFMMDPQSGKKICYV 598

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||:|||| |:
Db        599 QFPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAPKKKQ 658

Qy        650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAP 709
              || |||||||||| | || | || ||| | |  ::|    | :|     :||  |:|
Db        659 PPGRTCNCWPKWC-CLCC-GMRK-KKTGKVKDNQRK----KPKETSKQIHALEHIEEGLQ 711

Qy        710 GAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDK 769
                    |     | ||||||||| | | |||| |||   : :|||||:|:| |||||||:|
Db        712 VTNAENNSETAQLKLEKKFGQSPVLVASTLLLNGGVPSNVNPASLLRESIQVISCGYEEK 771

Qy        770 TDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWA 829
              |:|||||||||||||||||||||||||||||:||:||| |||||||:|||||||||||||
Db        772 TEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVLRWA 831

Qy        830 LGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI 889
              |||:||| | |||:|||||||||:|||||||||:||||||:||| ||:|||||||||||||
Db        832 LGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLLVYCSLPAICLLTGKFI 891

Qy        890 TPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGL 949
              ||::| | : |: :|: |  | ||||:|  :|||||||||||||||||||||||||||||
Db        892 VPEISNYAGILFLLMFMSIAVTGILEMQWGKIGIDDWWRNEQFWVIGGVSSHLFALFQGL 951

Qy        950 LKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNG 1009
              |||:||| |:|||||  || |||||| ||||:|||||||||::| :||: |||:|||||
Db        952 LKVLAGVSTNFTVTSKAADDGEFSELYIFKWTSLLIPPTTLLIINIVGVIVGVSDAINNG 1011

Qy       1010 YESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPF 1069
              |:|||||||:|||| ||||||||||||:|:|:| |||::|||||||| :|||||::||
Db       1012 YDSWGPLFGRLFFALWVIVHLYPFLKGLLGKQDRVPTIILVWSILLASILTLLWVRVNPF 1071

Qy       1070 LAKDDGPLLEECGLDC 1085
              ::| |||:|| |||||
Db       1072 VSK-DGPVLEICGLDC 1086
```

```
RESULT 13
ABB93411
ID   ABB93411 standard; Protein; 1069 AA.
XX
AC   ABB93411;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 2622.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
```

```
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 2622; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   1069 AA;

  Query Match             71.0%;  Score 4204.5;  DB 23;  Length 1069;
  Best Local Similarity   70.6%;  Pred. No. 0;
  Matches  770;  Conservative  126;  Mismatches  166;  Indels   29;  Gaps   10;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:||||||||  |:|  |      :   | :||  ||||||:: |:   |: ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVEELSGQTCQICGDEIELSVDGESFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              |||||||  |||||||||  |:||||||||||:||   ||  ||||:||:||||  ||::    |
Db         61 CAFPVCRPCYEYERREGNQSCPQCKTRYKRIKGSPRVEGDEEDDGIDDLDFEFDY----S 116

Qy        121 QSVAESMLYG--HMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMG 178
              :|  ||  :   :  :      | |:          :|||| |:   :|  : |||: |
Db        117 RSGLESETFSRRNSEFDLASAPPGS--------QIPLLTYGEEDVEISSDSHALIVSPSP 168

Qy        179 GGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQT---G 235
              |   |:|   : ||:    || | | |||| |||||||||:||| ||::|   :|
Db        169 GHIHRVHQPHFPDPA--AHPRPMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKYQVVKHD 226

Qy        236 NDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPV 295
              |    ||||||:|:||| || ||||:|: ||:||||||:|::|||:|| ||||||::|||
Db        227 GDSSLGDGDDADIPMMDEGRQPLSRKVPIKSSKINPYRMLIVLRLVILGLFFHYRILHPV 286

Qy        296 NDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFF 355
              |||:|||||||||||||:||:|||||||:|||||||||||||::|||:||:|| :| |
Db        287 NDAYALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSELAGVDVF 346

Qy        356 VSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVP 415
              ||||||:|||||:| |||||||:||||||:|:||||||||||||||||||:|||:||||
Db        347 VSTVDPMKEPPLITANTVLSILAVDYPVDRVACYVSDDGAAMLTFEALSETAEFARKWVP 406

Qy        416 FCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVP 475
              |||:| |||||||||  |:||||:||   ||||||||||:||||||:|||||| |||||
Db        407 FCKKYTIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQKVP 466

Qy        476 EEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKK 535
              |||||||||||||||||||||||||||||| :|  | | ||||||||||||||||::||||
Db        467 EEGWTMQDGTPWPGNNVRDHPGMIQVFLGNNGVRDVENNELPRLVYVSREKRPGFDHHKK 526

Qy        536 AGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRF 595
              |||||:|:||| ||:|||||||:||||||||||::|||||||  |||:|||||||||
Db        527 AGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRF 586

Qy        596 DGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTC 655
              ||||: |||:||||||||||||||:|||||||||||||||||:|||| ||    ||
Db        587 DGIDKSDRYSNRNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAPKKKKTKRMTC 646

Qy        656 NCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAEN-E 714
              |||||||  ||  | || :|:    |:|||    |  |     :||  |:|   |  :
Db        647 NCWPKWCLFCC--GLRKNRKS--KTTDKKK----KNREASKQIHALENIEEGTKGTNDAA 698

Qy        715 KAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGK 774
              |:    | ||||||||| ||| |  :|||| ::|||||||:||| ||||||||||:|||
Db        699 KSPEAAQLKLEKKFGQSPVFVASAGMENGGLARNASPASLLREAIQVISCGYEDKTEWGK 758

Qy        775 EIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIE 834
              |||||||||||||||||||| |||||:|| ||  |||||||:|||||||||||||||:|
Db        759 EIGWIYGSVTEDILTGFKMHSHGWRSVYCTPKIPAFKGSAPINLSDRLHQVLRWALGSVE 818

Qy        835 IFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELN 894
              || | |||:|||||||||:||| |||||:|||||||||| ||:|||||||||||  ||::
Db        819 IFLSRHCPIWYGYGGGLKWLERLSYINSVVYPWTSIPLLVYCSLPAICLLTGKFIVPEIS 878

Qy        895 NVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIA 954
              | ||: ||:||  |  | ||||:|  |||||||||||||||||:||||:|||||||:|
Db        879 NYASILFMALFGSIAVTGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLA 938

Qy        955 GVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWG 1014
              ||:|:|||||  || |||||| ||||:|||||||::| |||: |:|:||:|||:|||
Db        939 GVETNFTVTSKAADDGEFSELYIFKWTSLLIPPTTLLIINVIGVIVGISDAISNGYDSWG 998

Qy       1015 PLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDD 1074
              ||||:|||||||:|||||||||:|:|:| |||::|||||||| :|||||::||:||
Db        999 PLFGRLFFAFWVILHLYPFLKGLLGKQDRMPTIILVWSILLASILTLLWVRVNPFVAK-G 1057

Qy       1075 GPLLEECGLDC 1085
              ||:|| |||||
Db       1058 GPILEICGLDC 1068


RESULT 14
AAB37895
ID   AAB37895 standard; Protein; 1080 AA.
XX
AC   AAB37895;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Arabidopsis thaliana cellulose synthase CelA.
XX
KW   Arabidopsis thaliana; aspen; CelA; cellulose synthase; wood;
KW   mechanical stress response element; MSRE; cellulose biosynthesis.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200071670-A2.
XX
PD   30-NOV-2000.
XX
PF   18-MAY-2000; 2000WO-US13637.
XX
PR   21-MAY-1999;   99US-0135280.
XX
PA   (UNMT ) UNIV MICHIGAN TECHNOLOGICAL.
XX
PI   Chiang VL,  Wu L,  Chandrashekhar JP,  Carraway DT;
XX
DR   WPI; 2001-032025/04.
DR   N-PSDB; AAC83798.
XX
PT   Novel polynucleotide encoding cellulose synthase useful for enhancing
PT   cellulose biosynthesis, modifying lignin biosynthesis, and for
PT   accelerating growth of plants and for upregulating cellulose synthase
PT   levels  -
XX
PS   Disclosure; Fig 8; 65pp; English.
XX
CC   The present sequence is given in a specification relating to
CC   polynucleotides encoding cellulose synthase enzymes from Populus
CC   tremuloides and Arabidopsis. The cellulose synthases are designated
```

```
CC   PtCelA and AraxCelA respectively. The polynucleotides may be introduced
CC   into plants where they are useful for altering the growth of a plant.
CC   Vectors comprising a cellulose synthase promoter, or a fragment of the
CC   promoter containing one or more mechanical stress response elements
CC   (MSRE), operably linked to a cellulose synthase gene are useful for
CC   causing stress-induced gene expression in a plant cell. A cDNA encoding a
CC   protein that binds to a positive MSRE may be delivered into the plant to
CC   increase expression of cellulose synthase. A cDNA in an antisense
CC   orientation may be delivered to reduce expression of cellulose synthase.
CC   The polynucleotides enable the development of transgenic tree varieties
CC   having increased cellulose content, decreased lignin content and
CC   therefore improved wood fibre characteristics. The cellulose synthase
CC   promoters are useful for synthesising greater quantities of
CC   high-crystalline cellulose in plant and preferably in trees, which
CC   permits an increase in proportion of cellulose in transgenic plants,
CC   greater strength of juvenile wood and fiber and acceleration of overall
CC   growth rate.
CC   Note: This nucleotide sequence is given in the specification as
CC   SEQ ID NO: 5. SEQ ID NO: 5 is referred to in claim 8 but it is described
CC   as a polypeptide sequence.
XX
SQ   Sequence   1080 AA;

  Query Match            70.7%;  Score 4185.5;  DB 22;  Length 1080;
  Best Local Similarity  70.7%;  Pred. No. 0;
  Matches  769;  Conservative  126;  Mismatches  172;  Indels   21;  Gaps   11;

Qy          7 LVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVC 66
              |:|||||||| |:|  | :   :  :| :|| |||| |:: |   |:|||||||||||||
Db          4 LIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNECAFPVC 63

Qy         67 RDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW--DGHDSQSVA 124
              | |||||||| | |||||||:|||||  || |||||| :|||||||  :  :|     |:
Db         64 RPCYEYERREGNQACPQCKTRFKRLKGSPRVEGDEEEDDIDDLDNEFEYGNNGIGFDQVS 123

Qy        125 ESMLYGHMSYG-RGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL-VPSFMGGGGK 182
              | |     : |    | : ||   |    :|||| |    :|  ::||| ||  :|| |
Db        124 EGMSISRRNSGFPQSDLDSAPPGSQ----IPLLTYGDEDVEISSDRHALIVPPSLGGHGN 179

Qy        183 RIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERMHQTGNDGGGD 241
              |:||:  :||::    | | | |||| |||||||||:||| ||::| |::    ::|  |
Db        180 RVHPVSLSDPTVAAHRRLMVPQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVRHEGDPD 239

Qy        242 --DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAF 299
                |||||| |:||| || || ||||: ||:|||||:|::|||:|| |||||::||| ||:
Db        240 FEDGDDADFPMMDEGRQPLSMKIPIKSSKINPYRMLIVLRLVILGLFFHYRILHPVKDAY 299

Qy        300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTV 359
              |||||||||||||:||:|||||||:|||||||||||||::||||:|| |:|:| |||||
Db        300 ALWLISVICEIWFAVSWVLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVDVFVSTV 359

Qy        360 DPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 419
              ||||||||:| |||||||:||||||||:||||||||||||||||||:|||:|||||||:
Db        360 DPLKEPPLITANTVLSILAVDYPVDKVACYVSDDGAAMLTFEALSETAEFARKWVPFCKK 419

Qy        420 YNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 479
              | |||||||||||  |:||||:||   ||||||||||:||||||:|||||| |||||:||
Db        420 YCIEPRAPEWYFCHKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVATAQKVPEDGW 479

Qy        480 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAM 539
              ||||||||||:|||||||||||||  |  | | |||||||||||||||::||||||||
Db        480 TMQDGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPGFDHHKKAGAM 539

Qy        540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 599
              |:|:||| ||:|||||||:|||||||||||::|||||||||  |||:|||||||||||||
Db        540 NSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRFDGID 599

Qy        600 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 659
              |||||:|||||||||||||:|||||| ||||||||||:|||| || | :||||||
Db        600 RHDRYSNRNVVFFDINMKGLDGLQGPIYV-TGCVFRRQALYGFDAPKKKKGPRKTCNCWP 658

Qy        660 KWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA--PGAENEKAG 717
              |||    |||:|| :|      :|||    |  |     :||  |:|         |::
```

```
Db        659 KWCL--LCFGSRKNRKAKTVAADKKK----KNREASKQIHALENIEEGRGHKVLNVEQST 712

Qy        718 IVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIG 777
              | ||:||:||| ||| |  ||||  ::|||| |||||| ||| |||||||:||||||
Db        713 EAMQMKLQKKYGQSPVFVASARLENGGMARNASPACLLKEAIQVISRGYEDKTEWGKEIG 772

Qy        778 WIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFF 837
              |||||||||||| ||| |||| :|| ||  |||||||:||||||||||||||||:|||
Db        773 WIYGSVTEDILTGSKMHSHGWRHVYCTPKLAAFKGSAPINLSDRLHQVLRWALGSVEIFL 832

Qy        838 SNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVA 897
              | |||:|||||||||:||| |||||:||||||:||: ||:|||||||||||| ||::| |
Db        833 SRHCPIWYGYGGGLKWLERLSYINSVVYPWTSLPLIVYCSLPAICLLTGKFIVPEISNYA 892

Qy        898 SLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVD 957
              |: ||:||  |  | ||||:|  |||||||||||||||||:||||:|||||||:||||
Db        893 SILFMALFSSIAITGILEMQWGKVGIDDWWRNEQFWVIGGVSAHLFALFQGLLKVLAGVD 952

Qy        958 TSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLF 1017
              |:||||||  || |||:|| ||||:|||||  |||::| |||: |||:||:|||:||||||
Db        953 TNFTVTSKAADDGEFSDLYLFKWTSLLIPPMTLLIINVIGVIVGVSDAISNGYDSWGPLF 1012

Qy       1018 GKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPL 1077
              |:|||| |||:|||||||||:|:|:| |||::|||||||| :|||||::||:||  ||:
Db       1013 GRLPFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRVNPFVAK-GGPI 1071

Qy       1078 LEECGLDC 1085
              || |||||
Db       1072 LEICGLDC 1079


RESULT 15
AAY58837
ID   AAY58837 standard; Protein; 1039 AA.
XX
AC   AAY58837;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Soybean cellulose synthase (full-length).
XX
KW   Soybean; cellulose synthase; transgenic plant.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 201
FT                   /note= "encoded by AGAA (frameshift)"
XX
PN   WO200004166-A2.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15871.
XX
PR   14-JUL-1998;   98US-0092844.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Fader GM,  Falco SC,  Kinney AJ,  Lightner JE,  Miao G;
PI   Rafalski JA,  Thorpe CJ;
XX
DR   WPI; 2000-182431/16.
DR   N-PSDB; AAZ58268.
XX
PT   Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT   probes for isolating cDNAs and genes encoding homologous proteins, for
PT   producing transgenic plants  -
XX
PS   Claim 6; Page 56-60; 93pp; English.
XX
CC   The present sequence is that of full-length soybean cellulose
CC   synthase (CS), as deduced from an isolated cDNA clone (see AAY58268).
```

```
CC   The invention relates to isolated CS nucleic acid fragments and to
CC   CS polypeptides.  It also relates to the construction of a chimeric
CC   gene encoding all or a portion of the CS, in sense or antisense
CC   orientation, where expression of the gene results in altered levels
CC   of the CS in transformed host cells.  The host cells can be used to
CC   screen compounds for their ability to inhibit CS activity.  CS
CC   nucleic acids are also useful for producing transgenic plants having
CC   altered levels of CS, and hence altered levels of fibre.  CS may
CC   also serve as a target for the development of novel herbicides.
XX
SQ   Sequence   1039 AA;

  Query Match             67.4%;  Score 3986;  DB 21;  Length 1039;
  Best Local Similarity   67.6%;  Pred. No. 0;
  Matches  738;  Conservative  135;  Mismatches  159;  Indels   60;  Gaps   14;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||     |    ||  :  :||||:||||  |||    || ||||||
Db          1 MEASAGLVAGSHNRNELVVIH--GHEEPKALKNLDGQVCEICGDGVGLTVDGDLFVACNE 58

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              | |||||  |||||||||:   ||||||||||||    ||  ||::|: |:|:::|||  |    :
Db         59 CGFPVCRPCYEYERREGSHLCPQCKTRYKRLKGSPRVEGDDDEEDVDDIEHEFNIDEQKN 118

Qy        121 Q--SVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQ---MVDDIPPEQHALVPS 175
              :     |||:||:|  |||||  : :   |        |::   |:   :  :  |    :|
Db        119 KHGQVAEAMLHGRMSYGRGPEDDDNSQF-----PTPVIAGGRSRPVSGEFPISSNAYGDQ 173

Qy        176 FMGGG-GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQT 234
              :      ||:||  |  ::|         |   :   |                ||:||::||  :|
Db        174 MLSSSLHKRVHPYPVSEPG---SARWDEKKXD---------GWKDRMDDWKLQQ------ 215

Qy        235 GNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHP 294
              || |     | :|  |   ::|||||  ||||:|:  ||:||||||:|:  |||:|  ||    ||:|:|
Db        216 GNLGPEPD-EDPDAAMLDEARQPLSRKVPIASSKINPYRMVIVARLVILAFFLRYRLMNP 274

Qy        295 VNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDF 354
              |:||   |||  |:||||||| |||||||||||||:|||||||||:|:::||:|:  |||:|
Db        275 VHDALGLWLTSIICEIWFAFSWILDQFPKWFPIDRETYLDRLSIRYEREGEPNMLAPVDV 334

Qy        355 FVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWV 414
              |||||||:|||||||  ||||||::||||||:|||:||||||:|  |||:|||:|||:||||
Db        335 FVSTVDPMKEPPLVTANTVLSILAMDYPVDKISCYISDDGASMCTFESLSETAEFARKWV 394

Qy        415 PFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKV 474
              ||||:::|||||||| ||  :|||||||||    ||:||||||||||||||||||||||||||
Db        395 PFCKKFSIEPRAPEMYFSEKIDYLKDKVQPTFVKERRAMKREYEEFKVRINALVAKAQKV 454

Qy        475 PEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHK 534
              |:  ||  ||||||||||||  :||||||||||| |||||   |||:|||||||||||||:  |||
Db        455 PQGGWIMQDGTPWPGNNTKDHPGMIQVFLGSSGGLDTEGNQLPRLVYVSREKRPGFQHHK 514

Qy        535 KAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQR 594
              |||||||||||||||||||||::||||||||:||||||  :|||||:||||    |||||||||||
Db        515 KAGAMNALVRVSAVLTNAPFMLNLDCDHYVNNSKAAREAMCFLMDPQTGKKVCYVQFPQR 574

Qy        595 FDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRT 654
              ||||| ||||||||  ||||||||||||||||:|||||||||||||||||: ||    |  |
Db        575 FDGIDTHDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFRRQALYGYNPPKGPKRPKMV 634

Qy        655 CNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENE 714
                     | || |||:||:                ||:|       | ||     :   |  :::
Db        635 S------CDCCPCFGSRKK----------------YKEKND----ANGEA-ASLKGMDDD 667

Qy        715 KAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGK 774
              |   :::|    ||||||||:||||||:| ||    |:|||:||||||||||||||||||:||
Db        668 KEVLMSQMNFEKKFGQSSIFVTSTLMEEGGVPPSSSPAALLKEAIHVISCGYEDKTEWGL 727

Qy        775 EIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIE 834
              |:||||||:||||||||||||| ||||||||:|||  ||||:|:|:|||||||:||||||||||
Db        728 ELGWIYGSITEDILTGFKMHCRGWRSIYCMPKRAAFKGTAPINLSDRLNQVLRWALGSIE 787

Qy        835 IFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPEL 893
              ||||:|||||||:       ||:||||:| |: |||:|||||:|||  |||:||||  |||  |  :
Db        788 IFFSHHCPLWYGFKEKKLKWLERFAYANTTVYPFTSIPLVAYCILPAVCLLTDKFIMPPI 847

Qy        894 NNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVI 953
              :   |  |:|::||    |   ||  |||::||||  |::||||||||||||||:||||| ||||||:
Db        848 STFAGLYFVALFSSIIATGILELKWSGVSIEEWWRNEQFWVIGGVSAHLFAVIQGLLKVL 907

Qy        954 AGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESW 1013
              ||:||:||||||   |||||  |||||||||||||||||:|::|  :|||||:|:||||||:||
Db        908 AGIDTNFTVTSKATDDEEFGELYTFKWTTLLIPPTTILIINIVGVVAGISDAINNGYQSW 967

Qy       1014 GPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKD 1073
              ||||||||:|||||||||||||||:||||||||||::||:|||||||||||||||||: |
Db        968 GPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPTIVVIWSVLLASIFSLLWVRIDPFVLKT 1027

Qy       1074 DGPLLEECGLDC 1085
              ||   :  ||::|
Db       1028 KGPDTKLCGINC 1039

Search completed: December 15, 2003, 13:02:49
Job time : 51 secs
```

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:04:25 ; Search time 36 Seconds
(without alignments)
5610.507 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5918	100.0	1086	10	US-09-900-237-10	Sequence 10, Appl
2	5918	100.0	1086	12	US-10-209-059-50	Sequence 50, Appl
3	5918	100.0	1086	12	US-10-160-719-18	Sequence 18, Appl
4	5918	100.0	1086	12	US-10-160-719-38	Sequence 38, Appl
5	5570.5	94.1	1148	10	US-09-900-237-4	Sequence 4, Appli
6	5408.5	91.4	1059	12	US-10-209-059-42	Sequence 42, Appl
7	5408.5	91.4	1059	12	US-10-160-719-2	Sequence 2, Appli
8	5228	88.3	1094	12	US-10-209-059-18	Sequence 18, Appl
9	5228	88.3	1094	12	US-10-160-719-26	Sequence 26, Appl
10	5228	88.3	1094	12	US-10-160-719-46	Sequence 46, Appl
11	5228	88.3	1165	10	US-09-900-237-8	Sequence 8, Appli
12	4250.5	71.8	1084	15	US-10-229-193-8	Sequence 8, Appli
13	3991	67.4	1075	12	US-10-209-059-2	Sequence 2, Appli
14	3991	67.4	1075	12	US-10-160-719-10	Sequence 10, Appl
15	3991	67.4	1075	12	US-10-160-719-34	Sequence 34, Appl
16	3991	67.4	1075	12	US-10-160-719-54	Sequence 54, Appl
17	3986	67.4	1039	10	US-09-900-237-14	Sequence 14, Appl
18	3977.5	67.2	1074	12	US-10-209-059-46	Sequence 46, Appl
19	3977.5	67.2	1074	12	US-10-160-719-14	Sequence 14, Appl
20	3977.5	67.2	1074	12	US-10-160-719-22	Sequence 22, Appl
21	3977.5	67.2	1074	12	US-10-160-719-42	Sequence 42, Appl
22	3970.5	67.1	1091	10	US-09-900-237-26	Sequence 26, Appl
23	3945.5	66.7	1081	15	US-10-229-193-6	Sequence 6, Appli
24	3941.5	66.6	1081	15	US-10-229-193-12	Sequence 12, Appl
25	3917	66.2	1076	12	US-10-209-059-14	Sequence 14, Appl
26	3917	66.2	1076	12	US-10-160-719-58	Sequence 58, Appl
27	3905.5	66.0	1077	12	US-10-209-059-10	Sequence 10, Appl
28	3905.5	66.0	1077	12	US-10-160-719-6	Sequence 6, Appli
29	3905.5	66.0	1077	12	US-10-160-719-30	Sequence 30, Appl
30	3901	65.9	1078	12	US-10-209-059-26	Sequence 26, Appl
31	3880.5	65.6	1079	12	US-10-267-459-6	Sequence 6, Appli
32	3880.5	65.6	1079	12	US-10-209-059-22	Sequence 22, Appl
33	3876.5	65.5	1065	15	US-10-229-193-10	Sequence 10, Appl
34	3875.5	65.5	1065	10	US-09-900-237-33	Sequence 33, Appl
35	3818	64.5	1080	10	US-09-900-237-30	Sequence 30, Appl
36	3786.5	64.0	1043	12	US-10-160-719-50	Sequence 50, Appl
37	3611.5	61.0	927	12	US-10-209-059-30	Sequence 30, Appl
38	3552	60.0	881	10	US-09-838-539-8	Sequence 8, Appli
39	3521	59.5	793	10	US-09-900-237-18	Sequence 18, Appl
40	3469.5	58.6	974	10	US-09-838-539-6	Sequence 6, Appli
41	3442	58.2	1007	12	US-10-209-059-28	Sequence 28, Appl
42	3318	56.1	821	12	US-10-267-459-2	Sequence 2, Appli
43	3318	56.1	821	12	US-10-209-059-6	Sequence 6, Appli
44	3261.5	55.1	740	10	US-09-900-237-24	Sequence 24, Appl
45	2956.5	50.0	701	10	US-09-900-237-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-900-237-10
; Sequence 10, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
US-09-900-237-10

Query Match 100.0%; Score 5918; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy       1081 CGLDCN 1086
              ||||||
Db       1081 CGLDCN 1086
```

```
RESULT 2
US-10-209-059-50
; Sequence 50, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
;  APPLICANT: Dhugga, Kanwarpal S.
;  APPLICANT:  Wang, Haiyin
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864R2
;  CURRENT APPLICATION NUMBER: US/10/209,059
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 09/371,383
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: 09/550,483
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 1086
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-209-059-50

  Query Match             100.0%;  Score 5918;  DB 12;  Length 1086;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1086;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
```

```
Db        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Db        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Db        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Db        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Db        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Db       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy       1081 CGLDCN 1086

Db       1081 CGLDCN 1086
```

RESULT 3
US-10-160-719-18
; Sequence 18, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-18

Query Match 100.0%; Score 5918; DB 12; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db           1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy          61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db          61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy         121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Db         121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy         181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Db         181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy         241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHPVNDAFA 300

Db         241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy         301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Db         301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy         361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Db         361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy         421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Db         421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy         481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Db         481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy         541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Db         541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy         601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Db         601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy         661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Db         661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy         721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Db         721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy         781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Db         781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy         841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Db         841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy         901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Db         901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy         961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Db         961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy        1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Db        1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy        1081 CGLDCN 1086

Db        1081 CGLDCN 1086
```

RESULT 4
US-10-160-719-38
; Sequence 38, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof

```
;  FILE REFERENCE: 0864
;  CURRENT APPLICATION NUMBER: US/10/160,719
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/371,383A
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: US 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 38
;   LENGTH: 1086
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-160-719-38

  Query Match             100.0%;  Score 5918;  DB 12;  Length 1086;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1086;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGGG 180

Qy        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGGG 240

Qy        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFA 300

Qy        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVD 360

Qy        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRY 420

Qy        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWT 480

Qy        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMN 540

Qy        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDR 600

Qy        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPK 660

Qy        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 WCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVN 720

Qy        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 QQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIY 780

Qy        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNH 840

Qy        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLW 900

Qy        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 FMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSF 960

Qy        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKL 1020

Qy       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE 1080

Qy       1081 CGLDCN 1086
              ||||||
Db       1081 CGLDCN 1086


RESULT 5
US-09-900-237-4
; Sequence 4, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  TITLE OF INVENTION: Plant Cellulose Synthases
;  FILE REFERENCE: BB1170 US CIP
;  CURRENT APPLICATION NUMBER: US/09/900,237
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: 60/092,844
;  PRIOR FILING DATE: 1998-07-14
;  PRIOR APPLICATION NUMBER: PCT/US99/15871
;  PRIOR FILING DATE: 1999-07-13
;  PRIOR APPLICATION NUMBER: 09/720383
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
;   LENGTH: 1148
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-900-237-4

  Query Match             94.1%;  Score 5570.5;  DB 10;  Length 1148;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches 1027;  Conservative   23;  Mismatches   35;  Indels    5;  Gaps    3;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              ||||||||||||||||||||||:|||||  ::|||||||||||||  | |:|||||||
Db         60 MEASAGLVAGSHNRNELVVIRRDGEPGPKPMDQRNGQVCQICGDDVGRNPDGEPFVACNE 119

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNW-DGHD 119
              ||||:|||||||||||||||||||||:||||| || ||||||||||:||||| | ||
Db        120 CAFPICRDCYEYERREGTQNCPQCKTRFKRLKGCARVPGDEEEDGVDDLENEFNWSDKHD 179

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
              || :|||||: ||||||| | :| || |   |||||||||||||||:||||||||:||
Db        180 SQYLAESMLHAHMSYGRGADLDGVPQPFHPIPNVPLLTNGQMVDDIPPDQHALVPSFVGG 239

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQERMHQTGNDGG 239
              |||||||||||||:|||||||||||||||||||||||||||:|||:|||||||  ||||
Db        240 GGKRIHPLPYADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMESWKQKQERMHQTRNDGG 299

Qy        240 GDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAF 299
              |||||||||||||||| |||||||||||||||||||||||||| |||||||||| |||
Db        300 GDDGDDADLPLMDEARQPLSRKIPLPSSQINPYRMIIIIRLVVLCFFFHYRVMHPVPDAF 359

Qy        300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTV 359
              ||||||||||||||||||||||||||||||||||||||||||| ||||||:|||||||
```

```
Db        360 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGHPSQLAPVDFFVSTV 419

Qy        360 DPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 419
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 DPLKEPPLVTANTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKR 479

Qy        420 YNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 479
              |::|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        480 YSLEPRAPEWYFQQKIDYLKDKVAPNFVRERRAMKREYEEFKVRINALVAKAQKVPEEGW 539

Qy        480 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAM 539
              |||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db        540 TMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDVEGNELPRLVYVSREKRPGYNHHKKAGAM 599

Qy        540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID 659

Qy        600 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 RHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWP 719

Qy        660 KWCFCCCCFGNRKQK---KTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKA 716
              ||| |||||||| |   ||:||| || | | |||:|||:||||||||||||||||||||
Db        720 KWCICCCCFGNRKTKKKTKTSKPKFEKIKKL-FKKKENQAPAYALGEIDEAAPGAENEKA 778

Qy        717 GIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEI 776
              |||||||||||||||| |||||||||||||||||||||||||||||||||||| |||:|
Db        779 SIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYEDKTGWGKDI 838

Qy        777 GWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIF 836
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        839 GWIYGSVTEDILTGFKMHCHGWRSIYCIPKRAAFKGSAPLNLSDRLHQVLRWALGSIEIF 898

Qy        837 FSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNV 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 FSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNV 958

Qy        897 ASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGV 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 ASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGV 1018

Qy        957 DTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPL 1016
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db       1019 DTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINNGYESWGPL 1078

Qy       1017 FGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGP 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 FGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGP 1138

Qy       1077 LLEECGLDCN 1086
              ||||||||||
Db       1139 LLEECGLDCN 1148


RESULT 6
US-10-209-059-42
; Sequence 42, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
;  APPLICANT: Dhugga, Kanwarpal S.
;  APPLICANT:  Wang, Haiyin
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864R2
;  CURRENT APPLICATION NUMBER: US/10/209,059
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 09/371,383
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: 09/550,483
;  PRIOR FILING DATE: 2000-04-14

;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 1059
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-209-059-42

  Query Match             91.4%;  Score 5408.5;  DB 12;  Length 1059;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  995;  Conservative   22;  Mismatches   36;  Indels    5;  Gaps    3;

Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLK 92
              ::|||||||||||||  | |:|||||||||||:|||||||||||||||||||||:|| |
Db          3 QRNGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYEYERREGTQNCPQCKTRFKRFK 62

Qy         93 GCQRVTGDEEEDGVDDLDNEFNW-DGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNP 151
              || || ||||||||||||:||||| | |||| :|||||: ||||||| | :| || |   |
Db         63 GCARVPGDEEEDGVDDLENEFNWSDKHDSQYLAESMLHAHMSYGRGADLDGVPQPFHPIP 122

Qy        152 NVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYG 211
              |||||||||||||||:|||||||:||||||||||||||:||||||||||||||||||||
Db        123 NVPLLTNGQMVDDIPPDQHALVPSFVGGGGKRIHPLPYADPNLPVQPRSMDPSKDLAAYG 182

Qy        212 YGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINP 271
              |||||||||||:|||:|||||||| |||||||||||||||||||| ||||||||||||||
Db        183 YGSVAWKERMESWKQKQERMHQTRNDGGGDDGDDADLPLMDEARQPLSRKIPLPSSQINP 242

Qy        272 YRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 331
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db        243 YRMIIIIRLVVLCFFFHYRVMHPVPDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 302

Qy        332 YLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVS 391
              |||||||||||| ||||||:|||||||||||||||||| ||||||||||||||||||||
Db        303 YLDRLSLRFDKEGHPSQLAPVDFFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVSCYVS 362

Qy        392 DDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERR 451
              |||||||||||||||||||||||||||::|||||||||||||||||||| |||||||
Db        363 DDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQQKIDYLKDKVAPNFVRERR 422

Qy        452 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDC 511
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        423 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDV 482

Qy        512 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 542

Qy        572 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 602

Qy        632 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQK---KTTKPKTEKKKLLF 688
              |||||||||||||||||||||||||||||| ||||||||| |   ||:||| || | |
Db        603 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCICCCCFGNRKTKKKTKTSKPKFEKIKKL- 661

Qy        689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748
              |||:|||:|||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        662 FKKKENQAPAYALGEIDEAAPGAENEKASIVNQQKLEKKFGQSSVFVASTLLENGGTLKS 721

Qy        749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
              |||||||||||||||||||| |||:|||||||||||||||||||||||||||||||||||
Db        722 ASPASLLKEAIHVISCGYEDKTGWGKDIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRA 781

Qy        809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
              |||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||||
Db        782 AFKGSAPLNLSDRFHQVLRWALGSIEILFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 841

Qy        869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 901

Qy        929 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 988
```

```
Db        902 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 961

Qy        989 TLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1048
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TLLLLNFIGVVAGISNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1021

Qy       1049 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1086
              |||||||||||||||||||||||||||||||||||||
Db       1022 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1059


RESULT 7
US-10-160-719-2
; Sequence 2, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
;  APPLICANT: Kanwarpal S. Dhugga
;  APPLICANT:  Timothy G. Helentjaris
;  APPLICANT:  Benjamin A.  Bowen
;  APPLICANT:  Xun Wang
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864
;  CURRENT APPLICATION NUMBER: US/10/160,719
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/371,383A
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: US 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1059
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-160-719-2

  Query Match             91.4%;  Score 5408.5;  DB 12;  Length 1059;
  Best Local Similarity   94.0%;  Pred. No. 0;
  Matches  995;  Conservative   22;  Mismatches   36;  Indels    5;  Gaps    3;

Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLK 92
              ::|||||||||||||  | |:||||||||||:|||||||||||||||||||:|| |
Db          3 QRNGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYEYERREGTQNCPQCKTRFKRFK 62

Qy         93 GCQRVTGDEEEDGVDDLDNEFNW-DGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNP 151
              || || ||||||||||||:||||| | |||| :|||||: ||||||| | :| || |   |
Db         63 GCARVPGDEEEDGVDDLENEFNWSDKHDSQYLAESMLHAHMSYGRGADLDGVPQPFHPIP 122

Qy        152 NVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYG 211
              |||||||||||||||:||||||||:|||||||||||||||:||||||||||||||||||||
Db        123 NVPLLTNGQMVDDIPPDQHALVPSFVGGGGKRIHPLPYADPNLPVQPRSMDPSKDLAAYG 182

Qy        212 YGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINP 271
              |||||||||||:|||:|||||||  ||||||||||||||||||  ||||||||||||||
Db        183 YGSVAWKERMESWKQKQERMHQTRNDGGGDDGDDADLPLMDEARQPLSRKIPLPSSQINP 242

Qy        272 YRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 331
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||
Db        243 YRMIIIIRLVVLCFFFHYRVMHPVPDAFALWLISVICEIWFAMSWILDQFPKWFPIERET 302

Qy        332 YLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVS 391
              |||||||||||||  ||||||:||||||||||||||||  ||||||||||||||||||||
Db        303 YLDRLSLRFDKEGHPSQLAPVDFFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVSCYVS 362

Qy        392 DDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERR 451
              |||||||||||||||||||||||||::|||||||||||||||||||  |||||||
Db        363 DDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQQKIDYLKDKVAPNFVRERR 422

Qy        452 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDC 511
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        423 AMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGHDV 482
```

```
Qy        512 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 EGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIK 542

Qy        572 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 EAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTG 602

Qy        632 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQK---KTTKPKTEKKKLLF 688
              |||||||||||||||||||||||||||||| |||||||| |   ||:||| || | |
Db        603 CVFRRQALYGYDAPKTKKPPSRTCNCWPKWCICCCCFGNRKTKKKTKTSKPKFEKIKKL- 661
                                                                        o
Qy        689 FKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKS 748
              |||:|||:|||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db        662 FKKKENQAPAYALGEIDEAAPGAENEKASIVNQQKLEKKFGQSSVFVASTLLENGGTLKS 721

Qy        749 ASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRV 808
              |||||||||||||||||||| |||:|||||||||||||||||||||||||||||||||||
Db        722 ASPASLLKEAIHVISCGYEDKTGWGKDIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRA 781

Qy        809 AFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 868
              |||||||||||| |||||||||||| |||||||||||||||||||||||||||||||||||
Db        782 AFKGSAPLNLSDRFHQVLRWALGSIEILFSNHCPLWYGYGGGLKFLERFSYINSIVYPWT 841

Qy        869 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 928
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 SIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWR 901

Qy        929 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 988
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 NEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPT 961

Qy        989 TLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1048
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TLLLLNFIGVVAGISNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIV 1021

Qy       1049 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1086
              |||||||||||||||||||||||||||||||||||||
Db       1022 IVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDCN 1059


RESULT 8
US-10-209-059-18
; Sequence 18, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
;  APPLICANT: Dhugga, Kanwarpal S.
;  APPLICANT:  Wang, Haiyin
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864R2
;  CURRENT APPLICATION NUMBER: US/10/209,059
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 09/371,383
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: 09/550,483
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 1094
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-209-059-18

  Query Match             88.3%;  Score 5228;  DB 12;  Length 1094;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57
```

```
Db          1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116

Db         60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172

Db        120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229

Db        177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFH 288

Db        237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348

Db        297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408

Db        357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468

Db        417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528

Db        477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588

Db        537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596

Qy        589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648

Db        597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708

Db        657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768

Db        717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828

Db        777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888

Db        837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948

Db        897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008

Db        957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068

Db       1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy       1069 FLAKDDGPLLEECGLDCN 1086

Db       1077 FLAKSNGPLLEECGLDCN 1094
```

```
RESULT 9
US-10-160-719-26
; Sequence 26, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
;  APPLICANT: Kanwarpal S. Dhugga
;  APPLICANT:  Timothy G. Helentjaris
;  APPLICANT:  Benjamin A.  Bowen
;  APPLICANT:  Xun Wang
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864
;  CURRENT APPLICATION NUMBER: US/10/160,719
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/371,383A
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: US 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
;   LENGTH: 1094
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-160-719-26

  Query Match             88.3%;  Score 5228;  DB 12;  Length 1094;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57

Db          1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116

Db         60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172

Db        120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229

Db        177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFH 288

Db        237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348

Db        297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408

Db        357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468

Db        417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528

Db        477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588

Db        537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596
```

```
Qy       589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648
Db       597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy       649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
Db       657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy       709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
Db       717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy       769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
Db       777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy       829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
Db       837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy       889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
Db       897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy       949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
Db       957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy      1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
Db      1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy      1069 FLAKDDGPLLEECGLDCN 1086
Db      1077 FLAKSNGPLLEECGLDCN 1094
```

RESULT 10
US-10-160-719-46
; Sequence 46, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-46

Query Match 88.3%; Score 5228; DB 12; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

```
Qy         1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57
Db         1 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 59

Qy        58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116
```

```
Db        60 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 119

Qy       117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172
Db       120 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 176

Qy       173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229
Db       177 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 236

Qy       230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFH 288
Db       237 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 296

Qy       289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348
Db       297 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356

Qy       349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408
Db       357 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 416

Qy       409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468
Db       417 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 476

Qy       469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528
Db       477 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 536

Qy       529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588
Db       537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 596

Qy       589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648
Db       597 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 656

Qy       649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
Db       657 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 716

Qy       709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
Db       717 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 776

Qy       769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
Db       777 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 836

Qy       829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
Db       837 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 896

Qy       889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
Db       897 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 956

Qy       949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
Db       957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1016

Qy      1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
Db      1017 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1076

Qy      1069 FLAKDDGPLLEECGLDCN 1086
Db      1077 FLAKSNGPLLEECGLDCN 1094
```

RESULT 11
US-09-900-237-8

```
; Sequence 8, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  TITLE OF INVENTION: Plant Cellulose Synthases
;  FILE REFERENCE: BB1170 US CIP
;  CURRENT APPLICATION NUMBER: US/09/900,237
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: 60/092,844
;  PRIOR FILING DATE: 1998-07-14
;  PRIOR APPLICATION NUMBER: PCT/US99/15871
;  PRIOR FILING DATE: 1999-07-13
;  PRIOR APPLICATION NUMBER: 09/720383
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 1165
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-900-237-8

  Query Match             88.3%;  Score 5228;  DB 10;  Length 1165;
  Best Local Similarity   88.7%;  Pred. No. 0;
  Matches  974;  Conservative   38;  Mismatches   70;  Indels   16;  Gaps    7;

Qy          1 MEASAGLVAGSHNRNELVVIRRD---GDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVA 57
              |||||||||||||||||||||||    |  |     |       ||||||:||:   |:||||
Db         72 MEASAGLVAGSHNRNELVVIRRDRESGAAGGGAARRAEAP-CQICGDEVGVGFDGEPFVA 130

Qy         58 CNECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFN-WD 116
              |||||||||| |||||||||:| ||||:|||||||| || ||||||||||: ||    |
Db        131 CNECAFPVCRACYEYERREGSQACPQCRTRYKRLKGCPRVAGDEEEDGVDDLEGEFGLQD 190

Qy        117 G----HDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHAL 172
              |    | | ||||||   ||||||||   |   |   ||||||||||||||||||||||
Db        191 GAAHEDDPQYVAESMLRAQMSYGRGGD---AHPGFSPVPNVPLLTNGQMVDDIPPEQHAL 247

Qy        173 VPSFM---GGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQE 229
              |||:|   ||||||||||:|||:|||||||||||||||||||||||||||| |||:||
Db        248 VPSYMSGGGGGGKRIHPLPFADPNLPVQPRSMDPSKDLAAYGYGSVAWKERMEGWKQKQE 307

Qy        230 RMHQTGNDGGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFH 288
              |:    ::|||| |||||||||||||| ||||:|: ||:||||||||:||||||||||
Db        308 RLQHVRSEGGGDWDGDDADLPLMDEARQPLSRKVPISSSRINPYRMIIVIRLVVLGFFFH 367

Qy        289 YRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQ 348
              ||||||  |||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        368 YRVMHPAKDAFALWLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 427

Qy        349 LAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSE 408
              |||||||||||||| ||||||| |||||||||||||:|||||||||||||||||||||||
Db        428 LAPIDFFVSTVDPTKEPPLVTANTVLSILSVDYPVEKVSCYVSDDGAAMLTFEALSETSE 487

Qy        409 FAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALV 468
              |||||||| |::|||||||||||||||||||||||:|||||||||||||||||||||||
Db        488 FAKKWVPFSKKFNIEPRAPEWYFQQKIDYLKDKVAASFVRERRAMKREYEEFKVRINALV 547

Qy        469 AKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP 528
              |||||||||||||||:||||||||||||||||||| | ||||||||||||||||||||
Db        548 AKAQKVPEEGWTMQDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVYVSREKRP 607

Qy        529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCY 588
              |||||||||||||||||||:|| ||||||||||||||||||||||||||:||||||
Db        608 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKKVCY 667

Qy        589 VQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 648
              ||||||||||::|||||||||||||||||||||||||||||||||||||||||||||||
Db        668 VQFPQRFDGIDKNDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTK 727

Qy        649 KPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAA 708
              |||||||||||||  |||   |: :|||||||||||| ||||| || ||||||||||| |
Db        728 KPPSRTCNCWPKWCLSCCCSRNKNKKKTTKPKTEKKKRLFFKKAENPSPAYALGEIDEGA 787

Qy        709 PGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYED 768
              |||: |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        788 PGADIEKAGIVNQQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCGYED 847

Qy        769 KTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRW 828
              |||||||||||||:|||||||||||||||||||||| ||||||||||||||||||||||
Db        848 KTDWGKEIGWIYGSITEDILTGFKMHCHGWRSIYCIPKRPAFKGSAPLNLSDRLHQVLRW 907

Qy        829 ALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 888
              ||||:||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 ALGSVEIFFSKHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKF 967

Qy        889 ITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQG 948
              ||||| |||||:|||:|||||  | |||||||||| ||||||||||||||||:|||||||
Db        968 ITPELTNVASIWFMALFICISVTGILEMRWSGVAIDDWWRNEQFWVIGGVSAHLFAVFQG 1027

Qy        949 LLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINN 1008
              |||| ||:|||||||| ||||||||||||||||||||||||||||||||||:||||||
Db       1028 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGISNAINN 1087

Qy       1009 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDP 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       1088 GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRVDP 1147

Qy       1069 FLAKDDGPLLEECGLDCN 1086
              |||| :||||||||||||
Db       1148 FLAKSNGPLLEECGLDCN 1165


RESULT 12
US-10-229-193-8
; Sequence 8, Application US/10229193
; Publication No. US20030126643A1
; GENERAL INFORMATION:
;  APPLICANT: Arioli, Antonio Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION: Manipulation of Cellulose and/or Beta-1.4-Glucan
;  FILE REFERENCE: 96-98A
;  CURRENT APPLICATION NUMBER: US/10/229,193
;  CURRENT FILING DATE:  2002-08-26
;  PRIOR APPLICATION NUMBER: 09/221,013
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: PCT/AU97/00402
;  PRIOR FILING DATE: 1997-06-24
;  PRIOR APPLICATION NUMBER: AU PO0699
;  PRIOR FILING DATE: 1996-06-27
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 1084
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-229-193-8

  Query Match             71.8%;  Score 4250.5;  DB 15;  Length 1084;
  Best Local Similarity   70.8%;  Pred. No. 0;
  Matches  778;  Conservative   128;  Mismatches   162;  Indels   31;  Gaps   12;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |      |:|||||||| |:|  |      :  :| :|| ||||||:: |    : ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
              ||||||| ||||||||| | ||||||||||:||  || | ||||: :|||: ||:  | |
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEPD-HGMD 119

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
                :  ||: |   :: ||||  :  |        :||||       |:  ::|||:     |
Db        120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
```

```
Db        175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy        235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290

Db        235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy        291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350

Db        295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy        351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410

Db        355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSNNGAAMLTFEALSDTADFA 414

Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470

Db        415 TKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530

Db        475 AQKVPEERWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590

Db        535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650

Db        595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRKQALYGFDAPKKKKP 654

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKT--TKPKTEKKKLLFFKKEENQSPAYALGEIDEA- 707

Db        655 PGKTCNCWPKWC-CLCC-GLRKKSKTKATDKKTNTK--------ETSKQIHALENVDEGV 704

Qy        708 -APGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGY 766

Db        705 IVPVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGY 764

Qy        767 EDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVL 826

Db        765 QDKTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVL 824

Qy        827 RWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTG 886

Db        825 RWALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTG 884

Qy        887 KFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVF 946

Db        885 KFIVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALF 944

Qy        947 QGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAI 1006

Db        945 QGLLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAI 1004

Qy       1007 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRI 1066

Db       1005 SNGYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRI 1064

Qy       1067 DPFLAKDDGPLLEECGLDC 1085

Db       1065 NPFVAK-GGPVLEICGLNC 1082


RESULT 13
US-10-209-059-2
; Sequence 2, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
;  APPLICANT: Dhugga, Kanwarpal S.
;  APPLICANT:  Wang, Haiyin
;  TITLE OF INVENTION: Maize Cellulose Synthases and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0864R2
;  CURRENT APPLICATION NUMBER: US/10/209,059
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 09/371,383
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: 09/550,483
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1075
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-209-059-2

  Query Match             67.4%;  Score 3991;  DB 12;  Length 1075;
  Best Local Similarity   66.5%;  Pred. No. 0;
  Matches  739;  Conservative  143;  Mismatches  166;  Indels   64;  Gaps   15;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD-PG-PKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58

Db          1 MAANKGMVAGSHNRNEFVMIRHDGDVPGSAKPTKSANGQVCQICGDSVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118

Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGDEDEEDVDDLDNEFNY--- 117

Qy        119 DSQSVAESMLYGHMSYGRG------GDPNGAPQAFQLNPN--VPLLTNGQMVD----DIP 166

Db        118 ------------KQGSGKGPEWQLQGDDADLSSSARHEPHHRIPRLTSGQQISGEIPDAS 165

Qy        167 PEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQ 226

Db        166 PDRHSI----------RSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRV 215

Qy        227 RQER-MHQTGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMI 275

Db        216 KQDKNMMQVTNKYPEARGGDMEGTGSNGE--DMQMVDDARLPLSRIVPISSNQLNLYRVV 273

Qy        276 IIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDR 335

Db        274 IILRLIILCFFFQYRVSHPVRDAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDR 333

Qy        336 LSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGA 395

Db        334 LALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILSVDYPVDKVSCYVSDDGS 393

Qy        396 AMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKR 455

Db        394 AMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKR 453

Qy        456 EYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNE 515

Db        454 EYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNE 513

Qy        516 LPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMC 575

Db        514 LPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKALREAMC 573

Qy        576 FMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFR 635

Db        574 FMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFN 633

Qy        636 RQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQ 695

Db        634 RQALYGYD------PVLTEADLEPNIVIKSCC-GRRKKKNKSYMDSQSR---IMKRTESS 683

Qy        696 SPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755

Db        684 APIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPASLL 743
```

```
Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815
              |||||||||||||||:|||||||||||||||||||||||  ||:||||:| |  ||||||
Db        744 KEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFKGSAP 803

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875
              :||||||:|||||||||:||  | |||:|||| | || ||| :|||:|||| |||||:||
Db        804 INLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITSIPLIAY 863

Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935
              | |||||||| ||| ||::| | ::|: ||  |||| |||:|||||||:||||||||||
Db        864 CVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVI 923

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLN 994
              || |:|||||||||||:||:||:||||||  | | :|:||| ||||:||||||||:|::|
Db        924 GGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTVLVIN 983

Qy        995 FIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSIL 1054
              :|:|||:| |||:||:||||||||||: |||:|||||||||:||||||||||||||||
Db        984 LVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQNRTPTIVIVWSIL 1043

Qy       1055 LASIFSLLWVRIDPFLA-KDDGPLLEECGLDC 1085
              ||||||||||:||||::       | :||::|
Db       1044 LASIFSLLWVKIDPFISPTQKAAALGQCGVNC 1075
```

RESULT 14
US-10-160-719-10
; Sequence 10, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-10

Query Match 67.4%; Score 3991; DB 12; Length 1075;
Best Local Similarity 66.5%; Pred. No. 0;
Matches 739; Conservative 143; Mismatches 166; Indels 64; Gaps 15;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGD-PG-PKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
              | |: |:||||||||| |:|| ||| ||  || :  |||||||||| ||::  || ||||
Db          1 MAANKGMVAGSHNRNEFVMIRHDGDVPGSAKPTKSANGQVCQICGDSVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
              ||||||||| ||||||:|| | |||||||||| ||  || |||:|: |||||||||:
Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGDEDEEDVDDLDNEFNY--- 117

Qy        119 DSQSVAESMLYGHMSYGRG------GDPNGAPQAFQLNPN--VPLLTNGQMVD----DIP 166
                              |:|      ||       : :  |:  :| ||:|| :      |
Db        118 ------------KQGSGKGPEWQLQGDDADLSSSARHEPHHRIPRLTSGQQISGEIPDAS 165

Qy        167 PEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQ 226
              |::|::           |      | |||:||  | :|||||| :||  || ||||:|:|:
Db        166 PDRHSI----------RSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRV 215

Qy        227 RQER-MHQTGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMI 275
              :|:: | |  |          :| | :|:  |: ::|:||  ||| :|: |:|:| ||::
```

```
Db        216 KQDKNMMQVTNKYPEARGGDMEGTGSNGE--DMQMVDDARLPLSRIVPISSNQLNLYRVV 273

Qy        276 IIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDR 335
              ||:||::| ||| ||| ||| ||: |||:||||||:|||:||:||||||||:|| |||||||
Db        274 IILRLIILCFFFQYRVSHPVRDAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDR 333

Qy        336 LSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGA 395
              |:||:|:||:|||||||| |||||||||||||:| |||||||||||||||||||||||:
Db        334 LALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILSVDYPVDKVSCYVSDDGS 393

Qy        396 AMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKR 455
              |||||||:||||:|||:|||||||::|||||||:|| ||||||||||:  :||:|||||||
Db        394 AMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKR 453

Qy        456 EYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNE 515
              |||||||||||||||||||||||||| ||| ||||| |||||||||||| ||||| :|||
Db        454 EYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNE 513

Qy        516 LPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMC 575
              |||||||||||||: |||||||||||:|||||||  ||||:||||| |:|||::||||
Db        514 LPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKALREAMC 573

Qy        576 FMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFR 635
              ||||| ||:| |||||||||||| ||||||||:|||||||||||||||||:|||||| |
Db        574 FMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFN 633

Qy        636 RQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQ 695
              |||||||       |     :  |       || | ||:|  :   :: :      |: |:
Db        634 RQALYGYD------PVLTEADLEPNIVIKSCC-GRRKKKNKSYMDSQSR---IMKRTESS 683

Qy        696 SPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755
              :| : : :|:|    | |:|:: :::|:||||:|||| :|: || :  ||    | :|||||
Db        684 APIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPASLL 743

Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815
              |||||||||||||||:|||||||||||||||||||||||  ||:||||:| |  ||||||
Db        744 KEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFKGSAP 803

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875
              :||||||:|||||||||:||  | |||:|||| | || ||| :|||:|||| |||||:||
Db        804 INLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITSIPLIAY 863

Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935
              | |||||||| ||| ||::| | ::|: ||  |||| |||:|||||||:||||||||||
Db        864 CVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVI 923

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLN 994
              || |:|||||||||||:||:||:||||||  | | :|:||| ||||:||||||||:|::|
Db        924 GGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTVLVIN 983

Qy        995 FIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSIL 1054
              :|:|||:| |||:||:||||||||||: |||:|||||||||:||||||||||||||||
Db        984 LVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQNRTPTIVIVWSIL 1043

Qy       1055 LASIFSLLWVRIDPFLA-KDDGPLLEECGLDC 1085
              ||||||||||:||||::       | :||::|
Db       1044 LASIFSLLWVKIDPFISPTQKAAALGQCGVNC 1075
```

RESULT 15
US-10-160-719-34
; Sequence 34, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03

```
;  PRIOR APPLICATION NUMBER: US/09/371,383A
;  PRIOR FILING DATE: 1999-08-06
;  PRIOR APPLICATION NUMBER: US 60/096,822
;  PRIOR FILING DATE: 1998-08-17
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34
;   LENGTH: 1075
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-160-719-34

  Query Match             67.4%;  Score 3991;  DB 12;  Length 1075;
  Best Local Similarity   66.5%;  Pred. No. 0;
  Matches  739;  Conservative  143;  Mismatches  166;  Indels   64;  Gaps   15;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD-PG-PKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
              | |: |:|||||||||| |:|| ||| ||  || :  |||||||||| ||::  || ||||
Db          1 MAANKGMVAGSHNRNEFVMIRHDGDVPGSAKPTKSANGQVCQICGDSVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
              ||||||||| |||||||:|| | |||||||||| ||  || |||:|: |||||||||:
Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRQKGSPRVHGDEDEEDVDDLDNEFNY--- 117

Qy        119 DSQSVAESMLYGHMSYGRG------GDPNGAPQAFQLNPN--VPLLTNGQMVD----DIP 166
                          |:|      ||      : :  |:  :| ||:|| :     |
Db        118 ------------KQGSGKGPEWQLQGDDADLSSSARHEPHHRIPRLTSGQQISGEIPDAS 165

Qy        167 PEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQ 226
              |::|::          |     | |||:||  | :|||||| :||  || ||||:|:|:
Db        166 PDRHSI----------RSPTSSYVDPSVPVPVRIVDPSKDLNSYGLNSVDWKERVESWRV 215

Qy        227 RQER-MHQTGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMI 275
              :|:: | |  |          :| | :|:  |: ::|:||  ||| :|: |:|:| ||::
Db        216 KQDKNMMQVTNKYPEARGGDMEGTGSNGE--DMQMVDDARLPLSRIVPISSNQLNLYRVV 273

Qy        276 IIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDR 335
              ||:||::| ||| ||| ||| ||: |||:|||||:|||:||:|||||||:|| |||||||
Db        274 IILRLIILCFFFQYRVSHPVRDAYGLWLVSVICEVWFALSWLLDQFPKWYPINRETYLDR 333

Qy        336 LSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGA 395
              |:||:|:||:|||||||| |||||||||||:| |||||||||||||||||||||||||:
Db        334 LALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILSVDYPVDKVSCYVSDDGS 393

Qy        396 AMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKR 455
              ||||||:|||:|||:||||||::|||||||:|| |||||||||:  :||:|||||||
Db        394 AMLTFESLSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKR 453

Qy        456 EYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNE 515
              |||||||||||||||||||||||||| ||| ||||| |||||||||||| ||||| :|||
Db        454 EYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTDGNE 513

Qy        516 LPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMC 575
              |||||||||||||: |||||||||||:|||||||  ||||:||||| |:|||::||||
Db        514 LPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSSKALREAMC 573

Qy        576 FMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFR 635
              ||||| ||:| |||||||||||| ||||||||:|||||||||||||||:|||||| |
Db        574 FMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFN 633

Qy        636 RQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQ 695
              ||||||||      |     :  |       || | ||:|  :   :: :     |: |:
Db        634 RQALYGYD------PVLTEADLEPNIVIKSCC-GRRKKKNKSYMDSQSR---IMKRTESS 683

Qy        696 SPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755
              :| : : :|:|   | |:|:: :::|:||||:|||| :|: || :  ||   | :|||||
Db        684 APIFNMEDIEEGIEGYEDERSVLMSQRKLEKRFGQSPIFIASTFMTQGGIPPSTNPASLL 743

Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815
              |||||||||||||||:|||||||||||||||||||  ||:||||:| |  ||||||
Db        744 KEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWQSIYCMPPRPCFKGSAP 803

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875
              :||||||:|||||||||:||  | |||:|||| | || ||| :|||:|||| |||||:||
Db        804 INLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINTIVYPITSIPLIAY 863

Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935
              | |||||||| ||| ||::| | ::|: ||  |||| |||:|||||||:||||||||||
Db        864 CVLPAICLLTNKFIIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVI 923

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLN 994
              || |:|||||||||||:||:||:|||||  | | :|:|||| ||||:|||||||:|::|
Db        924 GGTSAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFAELYVFKWTSLLIPPTTVLVIN 983

Qy        995 FIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSIL 1054
               :|:|||:| |||:||:|||||||||||: |||:|||||||||:|||||||||||||||
Db        984 LVGMVAGISYAINSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGRQNRTPTIVIVWSIL 1043

Qy       1055 LASIFSLLWVRIDPFLA-KDDGPLLEECGLDC 1085
              ||||||||||:||||::       | :||::|
Db       1044 LASIFSLLWVKIDPFISPTQKAAALGQCGVNC 1075


Search completed: December 15, 2003, 13:10:07
Job time : 40 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:01:50 ; Search time 22 Seconds
(without alignments)
2088.618 Million cell updates/sec

Title: US-09-720-383C-10
Perfect score: 5918
Sequence: 1 MEASAGLVAGSHNRNELVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4250.5	71.8	1084	4	US-09-221-013A-8	Sequence 8, Appli
2	3945.5	66.7	1081	4	US-09-221-013A-6	Sequence 6, Appli
3	3941.5	66.6	1081	4	US-09-221-013A-12	Sequence 12, Appl
4	3876.5	65.5	1065	4	US-09-221-013A-10	Sequence 10, Appl
5	3552	60.0	881	3	US-08-960-048-8	Sequence 8, Appli
6	3552	60.0	881	4	US-09-838-586-8	Sequence 8, Appli
7	3469.5	58.6	974	3	US-08-960-048-6	Sequence 6, Appli
8	3469.5	58.6	974	4	US-09-838-586-6	Sequence 6, Appli
9	2902	49.0	685	3	US-08-960-048-7	Sequence 7, Appli
10	2902	49.0	685	4	US-09-838-586-7	Sequence 7, Appli
11	2436.5	41.2	629	4	US-09-221-013A-2	Sequence 2, Appli
12	1950.5	33.0	547	4	US-09-221-013A-14	Sequence 14, Appl
13	252	4.3	3031	1	US-07-689-008-2	Sequence 2, Appli
14	241.5	4.1	693	3	US-08-960-048-11	Sequence 11, Appl
15	241.5	4.1	693	4	US-09-838-586-11	Sequence 11, Appl
16	237	4.0	756	3	US-08-960-048-10	Sequence 10, Appl
17	237	4.0	756	4	US-09-147-236-3	Sequence 3, Appli
18	237	4.0	756	4	US-09-522-474-3	Sequence 3, Appli
19	237	4.0	756	4	US-09-838-586-10	Sequence 10, Appl
20	227.5	3.8	861	3	US-08-960-048-12	Sequence 12, Appl
21	227.5	3.8	861	4	US-09-838-586-12	Sequence 12, Appl
22	201	3.4	723	3	US-08-960-048-9	Sequence 9, Appli
23	201	3.4	723	4	US-08-541-939-2	Sequence 2, Appli
24	201	3.4	723	4	US-09-838-586-9	Sequence 9, Appli
25	201	3.4	723	5	PCT-US91-01726-4	Sequence 4, Appli
26	118.5	2.0	357	1	US-08-119-773-4	Sequence 4, Appli
27	117	2.0	1817	4	US-09-004-838-125	Sequence 125, App
28	110	1.9	1912	4	US-08-913-832A-2	Sequence 2, Appli
29	110	1.9	1912	4	US-09-249-181A-2	Sequence 2, Appli
30	110	1.9	1912	4	US-09-158-707-2	Sequence 2, Appli
31	109.5	1.9	357	1	US-08-119-773-6	Sequence 6, Appli
32	108.5	1.8	346	1	US-08-119-773-5	Sequence 5, Appli
33	108.5	1.8	357	1	US-08-119-773-2	Sequence 2, Appli
34	105	1.8	904	4	US-09-252-991A-19257	Sequence 19257, A
35	104	1.8	944	2	US-08-867-941-23	Sequence 23, Appl
36	104	1.8	944	3	US-09-074-658-23	Sequence 23, Appl
37	102	1.7	586	4	US-09-252-991A-28275	Sequence 28275, A
38	101.5	1.7	418	4	US-09-134-001C-4051	Sequence 4051, Ap
39	100	1.7	308	4	US-09-549-848B-37	Sequence 37, Appl
40	99.5	1.7	1416	1	US-08-061-465-4	Sequence 4, Appli
41	98.5	1.7	505	3	US-08-657-868B-4	Sequence 4, Appli
42	98.5	1.7	505	4	US-09-532-180A-4	Sequence 4, Appli
43	98	1.7	351	5	PCT-US91-06418-1	Sequence 1, Appli
44	98	1.7	501	4	US-09-328-352-5227	Sequence 5227, Ap
45	97.5	1.6	996	4	US-09-252-991A-27018	Sequence 27018, A

ALIGNMENTS

RESULT 1
US-09-221-013A-8
; Sequence 8, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-8

Query Match 71.8%; Score 4250.5; DB 4; Length 1084;
Best Local Similarity 70.8%; Pred. No. 0;
Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;

```
Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |     |:|||||||| |:|  |     :  :| :|| ||||||:: |    : ||||||
Db          1 MNTGGRLIAGSHNRNEFVLINADESARIRSVQELSGQTCQICGDEIELTVSSELFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTG-DEEEDGVDDLDNEFNWDGHD 119
              ||||||| |||||||| | |||||||||||:||   || | |||| : :|||: ||:  | |
Db         61 CAFPVCRPCYEYERREGNQACPQCKTRYKRIKGSPRVDGDDEEEEDIDDLEYEFD-HGMD 119

Qy        120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDDIPPEQHALVPSFMGG 179
              :  ||: |   :: |||| :  |        :||||       |:  ::|||:     |
Db        120 PEHAAEAALSSRLNTGRGGLDSAPP-----GSQIPLLTYCDEDADMYSDRHALIVPPSTG 174

Qy        180 GGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQ-ERM----HQT 234
              | |::| |: | | | | ||| | ||:| ||||||||||:||| ||:|| |::    |:
Db        175 YGNRVYPAPFTDSSAPPQARSMVPQKDIAEYGYGSVAWKDRMEVWKRRQGEKLQVIKHEG 234

Qy        235 GNDGGG----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYR 290
              ||:| |    |: || |:|:||| || ||||:|: ||:||||||:|: || :|| |||||
Db        235 GNNGRGSNDDDELDDPDMPMMDEGRQPLSRKLPIRSSRINPYRMLILCRLAILGLFFHYR 294

Qy        291 VMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLA 350
```

```
Db        295 ILHPVNDAYGLWLTSVICEIWFAVSWILDQFPKWYPIERETYLDRLSLRYEKEGKPSGLA 354

Qy        351 PIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFA 410

Db        355 PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVACYVSNNGAAMLTFEALSDTADFA 414

Qy        411 KKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470

Db        415 TKWVPFCKKFNIEPRAPEWYFSQKMDYLKNKVHPAFVRERRAMKRDYEEFKVKINALVAT 474

Qy        471 AQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGY 530

Db        475 AQKVPEERWTMQDGTPWPGNNVRDHPGMIQVFLGHSGVRDTDGNELPRLVYVSREKRPGF 534

Qy        531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQ 590

Db        535 DHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

Qy        591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKP 650

Db        595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPIYVGTGCVFRKQALYGFDAPKKKKP 654

Qy        651 PSRTCNCWPKWCFCCCCFGNRKQKKT--TKPKTEKKKLLFFKKEENQSPAYALGEIDEA- 707

Db        655 PGKTCNCWPKWC-CLCC-GLRKKSKTKATDKKTNTK--------ETSKQIHALENVDEGV 704

Qy        708 -APGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGY 766

Db        705 IVPVSNVEKRSEATQLKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCGY 764

Qy        767 EDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVL 826

Db        765 QDKTEWGKEIGWIYGSVTEDILTGFKMHCHGWRSVYCMPKRAAFKGSAPINLSDRLHQVL 824

Qy        827 RWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTG 886

Db        825 RWALGSVEIFLSRHCPIWYGYGGGLKWLERFSYINSVVYPWTSLPLIVYCSLPAVCLLTG 884

Qy        887 KFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVF 946

Db        885 KFIVPEISNYAGILFMLMFISIAVTGILEMQWGGVGIDDWWRNEQFWVIGGASSHLFALF 944

Qy        947 QGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAI 1006

Db        945 QGLLKVLAGVNTNFTVTSKAADDGAFSELYIFKWTTLLIPPTTLLIINIIGVIVGVSDAI 1004

Qy       1007 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRI 1066

Db       1005 SNGYDSWGPLFGRLFFALWVIVHLYPFLKGMLGKQDKMPTIIVVWSILLASILTLLWVRI 1064

Qy       1067 DPFLAKDDGPLLEECGLDC 1085

Db       1065 NPFVAK-GGPVLEICGLNC 1082


RESULT 2
US-09-221-013A-6
; Sequence 6, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 1081
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-221-013A-6

  Query Match             66.7%;  Score 3945.5;  DB 4;  Length 1081;
  Best Local Similarity   66.9%;  Pred. No. 0;
  Matches  727;  Conservative  140;  Mismatches  182;  Indels   37;  Gaps   11;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60

Db          1 MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMNGQICQICGDDVGLAETGDVFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120

Db         61 CAFPVCRPCYEYERKDGTQCCPQCKTRFRRHRGSPRVEGDEDEDDVDDIENEFNY----A 116

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDD---IPPEQHALVPSF- 176

Db        117 QGANKA------RHQRHGEEFSSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER--MHQT 234

Db        170 LGPSDRNAISSPYIDPRQPVPVRIVDPSKDLNSYGLGNVDWKERVEGWKLKQEKNMLQMT 229

Qy        235 GN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285

Db        230 GKYHEGKGGEIEGTGSNGE--ELQMADDTRLPMSRVVPIPSSRLTPYRVVIILRLIILCF 287

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDKEGQ 345

Db        288 FLQYRTTHPVKNAYPLWLTSVICEIWFAFSWLLDQFPKWYPINRETYLDRLAIRYDRDGE 347

Qy        346 PSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405

Db        348 PSQLVPVDVFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVACYVSDDGSAMLTFESLSE 407

Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465

Db        408 TAEFAKKWVPFCKKFNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRAMKREYEEFKVRIN 467

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525

Db        468 ALVAKAQKIPEEGWTMQDGTPWPGNNTRDHPGMIQVFLGHSGGLDTDGNELPRLIYVSRE 527

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585

Db        528 KRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK 587

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645

Db        588 CCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFNRQALYGYDPV 647

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705

Db        648 LTEE------DLEPNIIVKSCCGSRKKGKSSKKYNYEKRRGI--NRSDSNAPLFNMEDID 699

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765

Db        700 EGPEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIPPTTNPATLLKEAIHVISCG 759

Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825

Db        760 YEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCNPPRPAFKGSAPINLSDRLNQV 819

Qy        826 LRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLT 885

Db        820 LRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINTIVYPITSIPLIAYCILPAFCLIT 879

Qy        886 GKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAV 945
```

```
                :|| ||::| ||:||: ||| |  | |||:||||| |:|||||||||||||| |:|||||
Db        880 DRFIIPEISNYASIWFILLFISIAVTGILELRWSGVSIEDWWRNEQFWVIGGTSAHLFAV 939

Qy        946 FQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004
              |||||||:||:||:||||||  | | :|:||| |||| ||||||||:||:| ||:|||||
Db        940 FQGLLKVLAGIDTNFTVTSKATDEDGDFAELYIFKWTALLIPPTTVLLVNLIGIVAGVSY 999

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
              |:|:||:|||||||||||| ||| |||||||||:|||||||||||||||:||||||||||
Db       1000 AVNSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPTIVIVWSVLLASIFSLLWV 1059

Qy       1065 RIDPFL 1070
              ||:||:
Db       1060 RINPFV 1065


RESULT 3
US-09-221-013A-12
; Sequence 12, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 1081
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-221-013A-12

  Query Match           66.6%;  Score 3941.5;  DB 4;  Length 1081;
  Best Local Similarity  66.9%;  Pred. No. 0;
  Matches  726;  Conservative  140;  Mismatches  183;  Indels   37;  Gaps   11;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
              |||||||||||: ||||| || : | | || :  |||:||||||||||  || ||||||
Db          1 MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMNGQICQICGDDVGLAETGDVFVACNE 60

Qy         61 CAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGHDS 120
              ||||||| |||||||::||| |||||||::| :|  || |||:|| |||::||||:    :
Db         61 CAFPVCRPCYEYERKDGTQCCPQCKTRFRRHRGSPRVEGDEDEDDVDDIENEFNY----A 116

Qy        121 QSVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQMVDD---IPPEQHALVPSF- 176
              |   ::      : | |:   :    :  | :||||:|  |       |  |      |
Db        117 QGANKA------RHQRHGEEFSSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169

Qy        177 MGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER--MHQT 234
              :|   :     || ||  ||  | :|||||| :|| |:| ||||:| || :||:  :  |
Db        170 LGPSDRNAISSPYIDPRQPVPVRIVDPSKDLNSYGLGNVDWKERVEGWKLKQEKNMLQMT 229

Qy        235 GN---------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGF 285
              |         :| | :|:  :| : |: |  :|| :|:|:||:: |||::||:||::| |
Db        230 GKYHEGKGGEIEGTGSNGE--ELQMADDTRLPMSRVVPIPSSRLTPYRVVIILRLIILCF 287

Qy        286 FFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQ 345
              |  ||  ||| :|: ||| |||||||||| ||:|||||||:|| |||||||||::|:|::|:
Db        288 FLQYRTTHPVKNAYPLWLTSVICEIWFAFSWLLDQFPKWYPINRETYLDRLAIRYDRDGE 347

Qy        346 PSQLAPIDPFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSE 405
              |||| |:| |||||||||||||| |||||||||||||:|||||||:||||||:|||
Db        348 PSQLVPVDVFVSTVDPLKEPPLVTANTVLSILSVDYPVDKVACYVSDDGSAMLTFESLSE 407
```

```
Qy        406 TSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRIN 465
              |:|||||||||||::|||||||:|| |||||||||:  :||:|||||||||||||||||
Db        408 TAEFAKKWVPFCKKFNIEPRAPEYFYFAQKIDYLKDKIQPSFVKERRAMKREYEEFKVRIN 467

Qy        466 ALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525
              ||||||||:|||||||||||||||| |||||||||||| ||||| :|||||||:|||||
Db        468 ALVAKAQKIPEEGWTMQDGTPWPGNNTRDHPGMIQVFLGHSGGLDTDGNELPRLIYVSRE 527

Qy        526 KRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKK 585
              ||||: |||||||||||:||| |||| ||||:||||| |||||||||||||| :|||
Db        528 KRPGFQHHKKAGAMNALIRVSVVLTNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK 587

Qy        586 VCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP 645
              |||||||||||| ||||||||:|||||||||||||||:|||||| | ||||||||
Db        588 CCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGPVYVGTGCCFNRQALYGYDPV 647

Qy        646 KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEID 705
              |::      :  |       ||   :| | : |   ||:: :   : :: :| : : :||
Db        648 LTEE------DLEPNIIVKSCCGSRKKGKSSKKYNYEKRRGI--NRSDSNAPLFNMEDID 699

Qy        706 EAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCG 765
              |   | ::|:: :::|: :||:|||| ||: :| :| ||   : :||:||||||||||||
Db        700 EGFEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIPPTTNPATLLKEAIHVISCG 759

Qy        766 YEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQV 825
              |||||:|||||||||||||||||||||  || |||| | | |||||||:|||||:||
Db        760 YEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCNPPRPAFKGSAPINLSDRLNQV 819

Qy        826 LRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLT 885
              |||||||||  | |||:|||| | |: ||| :|||:|||| |||||:||| ||| ||:|
Db        820 LRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINTIVYPITSIPLIAYCILPAFCLIT 879

Qy        886 GKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAV 945
              :|| ||::| ||:||: ||| |  | |||:||||| |:|||||||||||| |:|||||
Db        880 DRFIIPEISNYASIWFILLFISIAVTGILELRWSGVSIEDWWRNEQFWVIGGTSAHLFAV 939

Qy        946 FQGLLKVIAGVDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSN 1004
              |||||||:||:||:||||||  | | :|:||| |||| ||||||||:||:| ||:|||||
Db        940 FQGLLKVLAGIDTNFTVTSKATDEDGDFAELYIFKWTALLIPPTTVLLVNLIGIVAGVSY 999

Qy       1005 AINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
              |:|:||:|||||||||||| ||| |||||||||:|||||||||||||||:||||||||||
Db       1000 AVNSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPTIVIVWSVLLASIFSLLWV 1059

Qy       1065 RIDPFL 1070
              ||:||:
Db       1060 RINPFV 1065


RESULT 4
US-09-221-013A-10
; Sequence 10, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 1065
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-221-013A-10
```

```
  Query Match             65.5%;  Score 3876.5;  DB 4;  Length 1065;
  Best Local Similarity   66.7%;  Pred. No. 0;
  Matches  728;  Conservative  145;  Mismatches  161;  Indels   57;  Gaps   18;

Qy         23 DGDPGPKPPREQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCP 82
              :|:   || :    | |||| |:||    || ||||: |:||||| |||||||:| |:||
Db          4 EGETAGKPMKNIVPQTCQICSDNVGKTVDGDRFVACDICSFPVCRPCYEYERKDGNQSCP 63

Qy         83 QCKTRYKRLKGCQRVTGDEEEDGV-DDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPN 141
              ||||||||||    : ||::|||: |:   |||:     : ::| ||   |:: |: |:
Db         64 QCKTRYKRLKGSPAIPGDKDEDGLADEGTVEFNYP--QKEKISERMLGWHLTRGK-GEEM 120

Qy        142 GAPQAFQ--LNPNVPLLTNGQMVDDIPPEQHALVPSFMG-----GGGKRIHPLPY-ADPS 193
              | ||  :   : ::| ||: |   |   |  |  |  :       ||||   ||| :| :
Db        121 GEPQYDKEVSHNHLPRLTSRQ---DTSGEFSAASPERLSVSSTIAGGKR---LPYSSDVN 174

Qy        194 LPVQPRSMDPSKDLAAYGYGSVAWKERMENWKQRQER-----MHQTGNDGGGDDGDD--- 245
                  | :||       | |:||||||:: || :||:       |  :: || | |
Db        175 QSPNRRIVDP------VGLGNVAWKERVDGWKMKQEKNTGPVSTQAASERGGVDIDASTD 228

Qy        246 --ADLPLM-DEARQQLSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALW 302
                ||  |: ||||| |||| :|||:||||||:|::|||:|  | |||: :|| :|||||
Db        229 ILADEALLNDEARQLLSRKVSIPSSRINPYRMVIMLRLVILCLFLHYRITNPVPNAFALW 288

Qy        303 LISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPL 362
              |:|||||||||:|||||||||||||: |||||||:||:|:||:|||| :| ||||||||
Db        289 LVSVICEIWFALSWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFVSTVDPL 348

Qy        363 KEPPLVTTNTVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNI 422
              ||||||| |||||||:||||||||||||||||||:||:|:||||||:|||||||:|:|
Db        349 KEPPLVTANTVLSILAVDYPVDKVSCYVSDDGAAMLSFESLAETSEFARKWVPFCKKYSI 408

Qy        423 EPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQ 482
              |||||||||  |||||||||  :||::||||||||||:||||||:|| | ||||| ||
Db        409 EPRAPEWYFAAKIDYLKDKVQTSFVKDRRAMKREYEEFKIRINALVSKALKCPEEGWVMQ 468

Qy        483 DGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNAL 542
              |||||||||  ||||||||||||:|||| ||||||||||||||||||: ||||||||||
Db        469 DGTPWPGNNTGDHPGMIQVFLGQNGGLDAEGNELPRLVYVSREKRPGFQHHKKAGAMNAL 528

Qy        543 VRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHD 602
              |||||||||  |::|||||||||||||||::|||||:||| |||:|||||||||||||::|
Db        529 VRVSAVLTNGPFILNLDCDHYINNSKALREAMCFLMDPNLGKQVCYVQFPQRFDGIDKND 588

Qy        603 RYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP---KTKKPP--SRTCNC 657
              ||||||  ||||||::||||||||:||||||| | |||||: |   | |||    |: |
Db        589 RYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPIKVKHKKPSLLSKLCG- 647

Qy        658 WPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGA--ENEK 715
                          |:||:     | :::|||     :  ::   | :  | :|:|    ||  ::||
Db        648 -----------GSRKKNSKAKKESDKKK--SGRHTDSTVPVFNLDDIEEGVEGAGFDDEK 694

Qy        716 AGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKE 775
              | :::|  |||:||||:||| |||:||||   ||:| : ||||||||||||||:||| |
Db        695 ALLMSQMSLEKRFGQSAVFVASTLMENGGVPPSATPENFLKEAIHVISCGYEDKSDWGME 754

Qy        776 IGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEI 835
              |||||||||||||||||  |||||||:||  |||||||:||||||:|||||||||:||
Db        755 IGWIYGSVTEDILTGFKMHARGWRSIYCMPKLPAFKGSAPINLSDRLNQVLRWALGSVEI 814

Qy        836 FFSNHCPLWYGYGGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNN 895
              || |||:|||| | |||||||:|:|: :|| |||||| |||| |:|| | :|| |:::|
Db        815 LFSRHCPIWYGYNGRLKFLERFAYVNTTIYPITSIPLLMYCTLLAVCLFTNQFIIPQISN 874

Qy        896 VASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAG 955
              :||:||:|||: |||| |||||||||||:|||||||||||||:|||||||:|||:||
Db        875 IASIWFLSLFLSIFATGILEMRWSGVGIDEWWRNEQFWVIGGVSAHLFAVFQGILKVLAG 934

Qy        956 VDTSFTVTSKGGD-DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWG 1014
              :||:||||||  | | :|:||| |||||||||||||::| :|||||| |||:||:|||
Db        935 IDTNFTVTSKASDEDGDFAELYLFKWTTLLIPPTTLLIVNLVGVVAGVSYAINSGYQSWG 994

Qy       1015 PLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDD 1074
              |||||||||||||||||||||:|||||||||:|||:|||||||||||||||||| ::
Db        995 PLFGKLFFAFWVIVHLYPFLKGLMGRQNRTPTIVVVWSVLLASIFSLLWVRIDPFTSRVT 1054

Qy       1075 GPLLEECGLDC 1085
              || : |||::|
Db       1055 GPDILECGINC 1065
```

```
RESULT 5
US-08-960-048-8
; Sequence 8, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
;  APPLICANT:  Stalker, D. et al.
;  TITLE OF INVENTION:  Plant Cellulose Synthase and Promoter
;  TITLE OF INVENTION:  Sequences
;  FILE REFERENCE:  15621/01/US
;  CURRENT APPLICATION NUMBER: US/08/960,048C
;  CURRENT FILING DATE:  1997-10-29
;  PRIOR APPLICATION NUMBER:  60/029,987
;  PRIOR FILING DATE:  1996-10-29
;  NUMBER OF SEQ ID NOS:  12
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 881
;   TYPE: PRT
;   ORGANISM: Oryzae sativa
US-08-960-048-8

  Query Match             60.0%;  Score 3552;  DB 3;  Length 881;
  Best Local Similarity   73.0%;  Pred. No. 0;
  Matches  651;  Conservative  103;  Mismatches  108;  Indels   30;  Gaps    7;

Qy        213 GSVAWKERMENWKQRQER----------MHQTGNDGGGD-----DGDDADLPLMDEARQQ 257
              |:||||||:: || :|::          :  :   | ||      | ::  |   | || ||
Db          1 GNVAWKERVDGWKLKQDKGAIPMTNGTSIAPSEGRGVGDIDASTDYNNEDALLNDETRQP 60

Qy        258 LSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWI 317
              ||||:|||||:|||| :|::|||||   | |||: :|| :|: |||:|||||||||:||:
Db         61 LSRKVPLPSSRINPYRNVIVLRLVVLSIFLHYRITNPVRNAYPLWLLSVICEIWFALSWL 120

Qy        318 LDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSIL 377
              :||||||||| ||||||||:||:|:||:|||| :| |||||||:|||||||  |||||||
Db        121 IDQFPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFVSTVDPMKEPPLVTANTVLSIL 180

Qy        378 SVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDY 437
              :|||||||||||||||||||:||:|||||||:|||||  |:|||||||||||  |||||
Db        181 AVDYPVDKVSCYVSDDGAAMLTFDALAETSEFARKWVPFVKKYNIEPRAPEWYFSQKIDY 240

Qy        438 LKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPG 497
              |||||  :||::|||||||||||||| ||||||||||| ||||||||||| |||||
Db        241 LKDKVHPSFVKDRRAMKREYEEFKVRINGLVAKAQKVPEEGWIMQDGTPWPGNNTRDHPG 300

Qy        498 MIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLN 557
              ||||||| ||||| |||||||||||||||||: |||||||||||||||||  |:||
Db        301 MIQVFLGHSGGLDTEGNELPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNGQYMLN 360

Qy        558 LDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMK 617
              |||||||||||::|||||:||| ||: |||||||||||||:||||||| ||||||::
Db        361 LDCDHYINNSKALREAMCFLMDPNLGRSVCYVQFPQRFDGIDRNDRYANRNTVFFDINLR 420

Qy        618 GLDGIQGPIYVGTGCVFRRQALYGYDAP-KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKT 676
              ||||||||:|||||||  | |||||: |  |  ||  |                | | :|  |:
Db        421 GLDGIQGPVYVGTGCVFNRTALYGYEPPIKQKKKGS---------FLSSLCGGRKKASKS 471

Qy        677 TKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGA--ENEKAGIVNQQKLEKKFGQSSVF 734
              |  ::|||    |  ::  | :  | :|:|    ||  ::||: :::|  |||:||||: |
Db        472 KKKSSDKKK--SNKHVDSAVPVFNLEDIEEGVEGAGFDDEKSLLMSQMSLEKRFGQSAAF 529

Qy        735 VTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMH 794
              | |||:| ||  :||:| |||||||||||||||||:|| ||||||||||||||||||
Db        530 VASTLMEYGGVPQSATPESLLKEAIHVISCGYEDKTEWGTEIGWIYGSVTEDILTGFKMH 589
```

```
Qy        795 CHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFL 854
              |||||||:||| |||||||:|||||:||||||||:|| || |||:|||||| ||||
Db        590 ARGWRSIYCMPKRPAFKGSAPINLSDRLNQVLRWALGSVEILFSRHCPIWYGYGGRLKFL 649

Qy        855 ERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSIL 914
              |||:|||: :|| |||||| || |||||||||||| ||::| ||:||:|||| |||| ||
Db        650 ERFAYINTTIYPLTSIPLLIYCVLPAICLLTGKFIIPEISNFASIWFISLFISIFATGIL 709

Qy        915 EMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFS 973
              ||||||||||:||||||||||||:|:||||||||||||:||:||:||||||  | | :|:
Db        710 EMRWSGVGIDEWWRNEQFWVIGGISAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFA 769

Qy        974 ELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPF 1033
              ||| |||||||||||:|::| :|||||:| |||:||:|||||||||||||||||||||||
Db        770 ELYMFKWTTLLIPPTTILIINLVGVVAGISYAINSGYQSWGPLFGKLFFAFWVIVHLYPF 829

Qy       1034 LKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDC 1085
              ||||:||||||||||:||:||||||||||||||  :  ||  : ||::|
Db        830 LKGLMGRQNRTPTIVVVWAILLASIFSLLWVRIDPFTTRVTGPDTQTCGINC 881


RESULT 6
US-09-838-586-8
; Sequence 8, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
;  APPLICANT: Stalker, D. et al.
;  TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
;  TITLE OF INVENTION:  Sequences
;  FILE REFERENCE: 15621/02/US
;  CURRENT APPLICATION NUMBER: US/09/838,586
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: 60/029,987
;  PRIOR FILING DATE: 1996-10-29
;  PRIOR APPLICATION NUMBER: 08/960,048
;  PRIOR FILING DATE: 1997-10-29
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 881
;   TYPE: PRT
;   ORGANISM: Oryzae sativa
US-09-838-586-8

  Query Match             60.0%;  Score 3552;  DB 4;  Length 881;
  Best Local Similarity   73.0%;  Pred. No. 0;
  Matches  651;  Conservative  103;  Mismatches  108;  Indels   30;  Gaps    7;

Qy        213 GSVAWKERMENWKQRQER----------MHQTGNDGGGD-----DGDDADLPLMDEARQQ 257
              |:||||||:: || :|::          :  :   | ||      | :: |  | || ||
Db          1 GNVAWKERVDGWKLKQDKGAIPMTNGTSIAPSEGRGVGDIDASTDYNNEDALLNDETRQP 60

Qy        258 LSRKIPLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWI 317
              ||||:|||||:||||| :|::|||||  | |||: :|| :|: |||:||||||||:||:
Db         61 LSRKVPLPSSRINPYRNVIVLRLVVLSIFLHYRITNPVRNAYPLWLLSVICEIWFALSWL 120

Qy        318 LDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSIL 377
              :||||||||| ||||||||:||:|:||:||||| :| |||||||:||||||| |||||||
Db        121 IDQFPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFVSTVDPMKEPPLVTANTVLSIL 180

Qy        378 SVDYPVDKVSCYVSDDGAAMLTPEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDY 437
              :||||||||||||||||||||:||:||||||:||||| |:|||||||||||| |||||
Db        181 AVDYPVDKVSCYVSDDGAAMLTFDALAETSEFARKWVPFVKKYNIEPRAPEWYFSQKIDY 240

Qy        438 LKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPG 497
              |||||  :||::|||||||||||||||| |||||||||||| |||||||||||| |||||
Db        241 LKDKVHPSFVKDRRAMKREYEEFKVRINGLVAKAQKVPEEGWIMQDGTPWPGNNTRDHPG 300

Qy        498 MIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLN 557
              ||||||| ||||| |||||||||||||||||: |||||||||||||||||||  |:||
Db        301 MIQVFLGHSGGLDTEGNELPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNGQYMLN 360
```

```
Qy        558 LDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMK 617
              ||||||||||||::|||||:||| ||: |||||||||||||||:||||||| ||||||::
Db        361 LDCDHYINNSKALREAMCFLMDPNLGRSVCYVQFPQRFDGIDRNDRYANRNTVFFDINLR 420

Qy        618 GLDGIQGPIYVGTGCVFRRQALYGYDAP-KTKKPPSRTCNCWPKWCFCCCCFGNRKQKKT 676
              ||||||||:|||||||| | |||||: | | ||  |                | | :|  |:
Db        421 GLDGIQGPVYVGTGCVFNRTALYGYEPPIKQKKKGS---------FLSSLCGGRKKASKS 471

Qy        677 TKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGA--ENEKAGIVNQQKLEKKFGQSSVF 734
              |  ::|||    |  ::  | : | :|:|   ||  ::||: :::|  |||:||||: |
Db        472 KKKSSDKKK--SNKHVDSAVPVFNLEDIEEGVEGAGFDDEKSLLMSQMSLEKRFGQSAAF 529

Qy        735 VTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMH 794
              | |||:| ||  :||:| ||||||||||||||||:|| ||||||||||||||||||
Db        530 VASTLMEYGGVPQSATPESLLKEAIHVISCGYEDKTEWGTEIGWIYGSVTEDILTGFKMH 589

Qy        795 CHGWRSIYCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFL 854
              |||||||:||| |||||||:|||||:||||||||:|| || |||:|||||| ||||
Db        590 ARGWRSIYCMPKRPAFKGSAPINLSDRLNQVLRWALGSVEILFSRHCPIWYGYGGRLKFL 649

Qy        855 ERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSIL 914
              |||:|||: :|| |||||| || |||||||||||| ||::| ||:||:|||| |||| ||
Db        650 ERFAYINTTIYPLTSIPLLIYCVLPAICLLTGKFIIPEISNFASIWFISLFISIFATGIL 709

Qy        915 EMRWSGVGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGD-DEEFS 973
              ||||||||||:||||||||||||:|:||||||||||||:||:||:||||||  | | :|:
Db        710 EMRWSGVGIDEWWRNEQFWVIGGISAHLFAVFQGLLKVLAGIDTNFTVTSKASDEDGDFA 769

Qy        974 ELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPF 1033
              ||| |||||||||||:|::| :|||||:| |||:||:|||||||||||||||||||||||
Db        770 ELYMFKWTTLLIPPTTILIINLVGVVAGISYAINSGYQSWGPLFGKLFFAFWVIVHLYPF 829

Qy       1034 LKGLVGRQNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEECGLDC 1085
              ||||:||||||||||:||:||||||||||||||  :  ||  : ||::|
Db        830 LKGLMGRQNRTPTIVVVWAILLASIFSLLWVRIDPFTTRVTGPDTQTCGINC 881


RESULT 7
US-08-960-048-6
; Sequence 6, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
;  APPLICANT:  Stalker, D. et al.
;  TITLE OF INVENTION:  Plant Cellulose Synthase and Promoter
;  TITLE OF INVENTION:  Sequences
;  FILE REFERENCE:  15621/01/US
;  CURRENT APPLICATION NUMBER: US/08/960,048C
;  CURRENT FILING DATE:  1997-10-29
;  PRIOR APPLICATION NUMBER:  60/029,987
;  PRIOR FILING DATE:  1996-10-29
;  NUMBER OF SEQ ID NOS:  12
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 974
;   TYPE: PRT
;   ORGANISM: Gossypim hirsutum
US-08-960-048-6

  Query Match             58.6%;  Score 3469.5;  DB 3;  Length 974;
  Best Local Similarity   59.8%;  Pred. No. 0;
  Matches  638;  Conservative  152;  Mismatches  168;  Indels  109;  Gaps   13;

Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKR-- 90
              |    ||  ||: |||   |:|||||:|| ||:|: |:||: :|| : | :| : |
Db          3 ESGVPVCHTCGEHVGLNVNGEPFVACHECNFPICKSCFEYDLKEGRKACLRCGSPYDENL 62

Qy         91 LKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLN 150
              |   :: |||                    :| :  |::  :
Db         63 LDDVEKATGD----------------------QSTMAAHLNKSQD------------ 85

Qy        151 PNVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAY 210
                                          :|   : |  :  |  :              |
Db         86 -------------------------VGIHARHISSVSTLDSEM-------------AE 105
```

```
Qy        211 GYGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLP----LMD----EARQQLSRKI 262
Db        106 DNGNSIWKNRVESWKEKKNKKKKPATT---KVEREAEIPPEQQMEDKPAPDASQPLSTII 162

Qy        263 PLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFP 322
Db        163 PIPKSRLAPYRTVIIMRLIILGLFFHYRVTNPVDSAFGLWLTSVICEIWFAFSWVLDQFP 222

Qy        323 KWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYP 382
Db        223 KWYPVNRETYIDRLSARYEREGEPDELAAVDFFVSTVDPLKEPPLITANTVLSILALDYP 282

Qy        383 VDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKV 442
Db        283 VDKVSCYISDDGAAMLTFESLVETADFARKWVPFCKKFSIEPRAPEFYFSQKIDYLKDKV 342

Qy        443 AANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVF 502
Db        343 QPSFVKERRAMKRDYEEYKIRINALVAKAQKTPDEGWTMQDGTSWPGNNPRDHPGMIQVF 402

Qy        503 LGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDH 562
Db        403 LGYSGARDIEGNELPRLVYVSREKRPGYQHHKKAGAENALVRVSAVLTNAPFILNLDCDH 462

Qy        563 YINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGI 622
Db        463 YVNNSKAVREAMCFLMDPQVGRDVCYVQFPQRFDGIDRSDRYANRNTVFFDVNMKGLDGI 522

Qy        623 QGPIYVGTGCVFRRQALYGYDAPKTKK-PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKT 681
Db        523 QGPVYVGTGCVFNRQALYGYGPPSMPSFPKSSSSSC------SCCCPGKK------EPKD 570

Qy        682 EKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLE 741
Db        571 PSELYRDAKREELDAAIFNLREIDNY---DEYERSMLISQTSFEKTFGLSSVFIESTLME 627

Qy        742 NGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSI 801
Db        628 NGGVAESANPSTLIKEAIHVISCGYEEKTAWGKEIGWIYGSVTEDILTGFKMHCRGWRSI 687

Qy        802 YCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG-LKFLERFSYI 860
Db        688 YCMPLRPAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPLWYGFGGGRLKWLQRLAYI 747

Qy        861 NSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSG 920
Db        748 NTIVYPFTSLPLIAYCSLPAICLLTGKFIIPTLSNLASVLFLGLFLSIIVTAVLELRWSG 807

Qy        921 VGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKW 980
Db        808 VSIEDLWRNEQFWVIGGVSAHLFAVFQGFLKMLAGIDTNFTVTAKAADDADFGELYIVKW 867

Qy        981 TTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGR 1040
Db        868 TTLLIPPTTLLIVNMVGVVAGFSDALNKGYEAWGPLFGKVFFSFWVILHLYPFLKGLMGR 927

Qy       1041 QNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE--CGLDC 1085
Db        928 QNRTPTIVVLWSVLLASVFSLVWVRINPFVSTADSTTVSQSCISIDC 974
```

RESULT 8
US-09-838-586-6
; Sequence 6, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypim hirsutum
US-09-838-586-6

Query Match 58.6%; Score 3469.5; DB 4; Length 974;
Best Local Similarity 59.8%; Pred. No. 0;
Matches 638; Conservative 152; Mismatches 168; Indels 109; Gaps 13;

```
Qy         33 EQNGQVCQICGDDVGLAPGGDPFVACNECAFPVCRDCYEYERREGTQNCPQCKTRYKR-- 90
Db          3 ESGVPVCHTCGEHVGLNVNGEPFVACHECNFPICKSCFEYDLKEGRKACLRCGSPYDENL 62

Qy         91 LKGCQRVTGDEEEDGVDDLDNEFNWDGHDSQSVAESMLYGHMSYGRGGDPNGAPQAFQLN 150
Db         63 LDDVEKATGD----------------------QSTMAAHLNKSQD------------- 85

Qy        151 PNVPLLTNGQMVDDIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAY 210
Db         86 -------------------------VGIHARHISSVSTLDSEM-------------AE 105

Qy        211 GYGSVAWKERMENWKQRQERMHQTGNDGGGDDGDDADLP----LMD----EARQQLSRKI 262
Db        106 DNGNSIWKNRVESWKEKKNKKKKPATT---KVEREAEIPPEQQMEDKPAPDASQPLSTII 162

Qy        263 PLPSSQINPYRMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFP 322
Db        163 PIPKSRLAPYRTVIIMRLIILGLFFHYRVTNPVDSAFGLWLTSVICEIWFAFSWVLDQFP 222

Qy        323 KWFPIERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYP 382
Db        223 KWYPVNRETYIDRLSARYEREGEPDELAAVDFFVSTVDPLKEPPLITANTVLSILALDYP 282

Qy        383 VDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKV 442
Db        283 VDKVSCYISDDGAAMLTFESLVETADFARKWVPFCKKFSIEPRAPEFYFSQKIDYLKDKV 342

Qy        443 AANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVF 502
Db        343 QPSFVKERRAMKRDYEEYKIRINALVAKAQKTPDEGWTMQDGTSWPGNNPRDHPGMIQVF 402

Qy        503 LGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDH 562
Db        403 LGYSGARDIEGNELPRLVYVSREKRPGYQHHKKAGAENALVRVSAVLTNAPFILNLDCDH 462

Qy        563 YINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGI 622
Db        463 YVNNSKAVREAMCFLMDPQVGRDVCYVQFPQRFDGIDRSDRYANRNTVFFDVNMKGLDGI 522

Qy        623 QGPIYVGTGCVFRRQALYGYDAPKTKK-PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKT 681
Db        523 QGPVYVGTGCVFNRQALYGYGPPSMPSFPKSSSSSC------SCCCPGKK------EPKD 570

Qy        682 EKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLE 741
Db        571 PSELYRDAKREELDAAIFNLREIDNY---DEYERSMLISQTSFEKTFGLSSVFIESTLME 627

Qy        742 NGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSI 801
Db        628 NGGVAESANPSTLIKEAIHVISCGYEEKTAWGKEIGWIYGSVTEDILTGFKMHCRGWRSI 687

Qy        802 YCIPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG-LKFLERFSYI 860
Db        688 YCMPLRPAFKGSAPINLSDRLHQVLRWALGSVEIFLSRHCPLWYGFGGGRLKWLQRLAYI 747

Qy        861 NSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSG 920
```

```
Db        748 NTIVYPPTSLPLIAYCSLPAICLLTGKFIIPTLSNLASVLFLGLFLSIIVTAVLELRWSG 807

Qy        921 VGIDDWWRNEQFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKW 980

Db        808 VSIEDLWRNEQFWVIGGVSAHLFAVFQGFLKMLAGIDTNFTVTAKAADDADFGELYIVKW 867

Qy        981 TTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGR 1040

Db        868 TTLLIPPTTLLIVNMVGVVAGFSDALNKGYEAWGPLFGKVFFSFWVILHLYPFLKGLMGR 927

Qy       1041 QNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDDGPLLEE--CGLDC 1085

Db        928 QNRTPTIVVLWSVLLASVFSLVWVRINPFVSTADSTTVSQSCISIDC 974
```

RESULT 9
US-08-960-048-7
; Sequence 7, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-7

Query Match 49.0%; Score 2902; DB 3; Length 685;
Best Local Similarity 75.0%; Pred. No. 9.9e-284;
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

```
Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469

Db          1 ARRWVPFCKKHNVEPRAPEFYFNEKIDYLKDKVHPSFVKERRAMKREYEEFKVRINALVA 60

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529

Db         61 KAQKKPEEGWVMQDGTPWPGNNTRDHPGMIQVYLGSAGALDVDGKELPRLVYVSREKRPG 120

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589

Db        121 YQHHKKAGAENALVRVSAVLTNAPFILNLDCDHYINNSKAMREAMCFLMDPQFGKKLCYV 180

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649

Db        181 QFPQRFDGIDRHDRYANRNVVFFDINMLGLDGLQGPVYVGTGCVFNRQALYGYDPPVSEK 240

Qy        650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKK----KLLFFKK---------EENQS 696

Db        241 RPKMTCDCWPSWC-CCCCGGSRKK---SKKKGEKKGLLGGLLYGKKKKMMGKNYVKKGSA 296

Qy        697 PAYALGEIDEAAPGAEN-EKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755

Db        297 PVFDLEEIEEGLEGYEELEKSTLMSQKNFEKRFGQSPVFIASTLMENGGLPEGTNSTSLI 356

Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815

Db        357 KEAIHVISCGYEEKTEWGKEIGWIYGSVTEDILTGFKMHCRGWKSVYCVPKRPAFKGSAP 416

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875

Db        417 INLSDRLHQVLRWALGSVEIFLSRHCPLWYGYGGKLKWLERLAYINTIVYPFTSIPLLAY 476
```

```
Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935

Db        477 CTIPAVCLLTGKFIIPTLSNLTSVWFLALFLSIIATGVLELRWSGVSIQDWWRNEQFWVI 536

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNF 995

Db        537 GGVSAHLFAVFQGLLKVLAGVDTNFTVTAKAADDTEFGELYLFKWTTLLIPPTTLIILNM 596

Qy        996 IGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILL 1055

Db        597 VGVVAGVSDAINNGYGSWGPLFGKLFFAFWVILHLYPFLKGLMGRQNRTPTIVVLWSILL 656

Qy       1056 ASIFSLLWVRIDPFLAKDDGPLLEECGLD 1084

Db        657 ASIFSLVWVRIDPFLPKQTGPVLKQCGVE 685
```

RESULT 10
US-09-838-586-7
; Sequence 7, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-586-7

Query Match 49.0%; Score 2902; DB 4; Length 685;
Best Local Similarity 75.0%; Pred. No. 9.9e-284;
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

```
Qy        410 AKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVA 469

Db          1 ARRWVPFCKKHNVEPRAPEFYFNEKIDYLKDKVHPSFVKERRAMKREYEEFKVRINALVA 60

Qy        470 KAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPG 529

Db         61 KAQKKPEEGWVMQDGTPWPGNNTRDHPGMIQVYLGSAGALDVDGKELPRLVYVSREKRPG 120

Qy        530 YNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYV 589

Db        121 YQHHKKAGAENALVRVSAVLTNAPFILNLDCDHYINNSKAMREAMCFLMDPQFGKKLCYV 180

Qy        590 QFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKK 649

Db        181 QFPQRFDGIDRHDRYANRNVVFFDINMLGLDGLQGPVYVGTGCVFNRQALYGYDPPVSEK 240

Qy        650 PPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKK----KLLFFKK---------EENQS 696

Db        241 RPKMTCDCWPSWC-CCCCGGSRKK---SKKKGEKKGLLGGLLYGKKKKMMGKNYVKKGSA 296

Qy        697 PAYALGEIDEAAPGAEN-EKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLL 755

Db        297 PVFDLEEIEEGLEGYEELEKSTLMSQKNFEKRFGQSPVFIASTLMENGGLPEGTNSTSLI 356

Qy        756 KEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAFKGSAP 815

Db        357 KEAIHVISCGYEEKTEWGKEIGWIYGSVTEDILTGFKMHCRGWKSVYCVPKRPAFKGSAP 416

Qy        816 LNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSIPLLAY 875
```

```
Db        417 INLSDRLHQVLRWALGSVEIFLSRHCPLWYGYGGKLKWLERLAYINTIVYPFTSIPLLAY 476

Qy        876 CTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNEQFWVI 935
Db        477 CTIPAVCLLTGKFIIPTLSNLTSVWFLALFLSIIATGVLELRWSGVSIQDWWRNEQFWVI 536

Qy        936 GGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGGDDEEFSELYTFKWTTLLIPPTTLLLLNF 995
Db        537 GGVSAHLFAVFQGLLKVLAGVDTNFTVTAKAADDTEFGELYLFKWTTLLIPPTTLIILNM 596

Qy        996 IGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWSILL 1055
Db        597 VGVVAGVSDAINNGYGSWGPLFGKLFFAFWVILHLYPFLKGLMGRQNRTPTIVVLWSILL 656

Qy       1056 ASIFSLLWVRIDPFLAKDDGPLLEECGLD 1084
Db        657 ASIFSLVWVRIDPFLPKQTGPVLKQCGVE 685


RESULT 11
US-09-221-013A-2
; Sequence 2, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 629
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-221-013A-2

  Query Match             41.2%;  Score 2436.5;  DB 4;  Length 629;
  Best Local Similarity   71.0%;  Pred. No. 9.3e-237;
  Matches  441;  Conservative   78;  Mismatches   93;  Indels    9;  Gaps    3;

Qy        451 RAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLD 510
Db          1 RAMKREYEEFKVRINALVAKAQKIPGEGWTMQDGTPWPGNNTRDHPGMIQVFLGHSGGLD 60

Qy        511 CEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAI 570
Db         61 TDGNELPRLIYVSREKRPGFQHHKKAGAMNASIRVSAVLTNGAYLLNVDCDHYFNNSKAI 120

Qy        571 KEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPIYVGT 630
Db        121 KEAMCFMMDPAIGKKCCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIHGPVYVGT 180

Qy        631 GCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPKTEKKKLLFFK 690
Db        181 GCCFNRQALYGYDPVLTEE------DLEPNIIVKSCCGSRKKGKSSKKYNYEKRRGI--N 232

Qy        691 KEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLLENGGTLKSAS 750
Db        233 RSDSNAPLFNMEDIDEGFEGYDDERSILMSQRSVEKRFGQSPVFIAATFMEQGGIPPTTN 292

Qy        751 PASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGFKMHCHGWRSIYCIPKRVAF 810
Db        293 PATLLKEAIHVISCGYEDKTEWGKEIGWIYGSVTEDILTGFKMHARGWISIYCNPPRPAF 352
```

```
Qy        811 KGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFSYINSIVYPWTSI 870
Db        353 KGSAPINLSDRLNQVLRWALGSIEILLSRHCPIWYGYHGRLRLLERIAYINTIVYPITSI 412

Qy        871 PLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFATSILEMRWSGVGIDDWWRNE 930
Db        413 PLIAYCILPAFCLITDRFIIPEISNYASIWFILLFISIAVTGILKLKWNGVSIEDWWRNN 472

Qy        931 QFWVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKG-GDDEEFSELYTFKWTTLLIPPTT 989
Db        473 QFWVIGGTSTHLFAVFQGLLKVLAGINTNFTVTSKATNKNGDFAKLYIFKWTALLIPPTT 532

Qy        990 LLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVI 1049
Db        533 VLLVNLIGIVAGVSYAVNSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPTIVI 592

Qy       1050 VWSILLASIFSLLWVRIDPFL 1070
Db        593 VWSVLLASIFSLLWVRINPFV 613


RESULT 12
US-09-221-013A-14
; Sequence 14, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 547
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-09-221-013A-14

  Query Match             33.0%;  Score 1950.5;  DB 4;  Length 547;
  Best Local Similarity   64.1%;  Pred. No. 9.4e-188;
  Matches  369;  Conservative   66;  Mismatches   84;  Indels   57;  Gaps    8;

Qy          1 MEASAGLVAGSHNRNELVVIRRDGD--PGPKPPREQNGQVCQICGDDVGLAPGGDPFVAC 58
Db          1 MAANAGMVAGSRNRNEFVMIRPDGDAPPPAKPGKSVNGQVCQICGDTVGVSATGDVFVAC 60

Qy         59 NECAFPVCRDCYEYERREGTQNCPQCKTRYKRLKGCQRVTGDEEEDGVDDLDNEFNWDGH 118
Db         61 NECAFPVCRPCYEYERKEGNQCCPQCKTRYKRHKGCPRVQGDEEEEDVDDLDNEF----- 115

Qy        119 DSQSVAESMLYGHMSYGRGGDPNGAPQ-----------AFQLNPNVPLLTNGQMVD---- 163
Db        116 -----------HYKHGNGKGPEWQIQRQGEDVDLSSSSRHEQHRIPRLTSGQQISGEIP 163

Qy        164 DIPPEQHALVPSFMGGGGKRIHPLPYADPSLPVQPRSMDPSKDLAAYGYGSVAWKERMEN 223
Db        164 DASPDRHSI----------RSGTSSYVDPSVPVPVRIVDPSKDLNSYGINSVDWQERVAS 213

Qy        224 WKQRQER-MHQTGN----------DGGGDDGDDADLPLMDEARQQLSRKIPLPSSQINPY 272
Db        214 WRNKQDKNMMQVANKYPEARGGDMEGTGSNGE--DIQMVDDARLPLSRIVPIPSNQLNLY 271

Qy        273 RMIIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWFPIERETY 332
Db        272 RIVIILRLIILMFFFQYRVTHPVRDAYGLWLVSVICEIWLPLSWLLDQFPKWYPINRETY 331
```

```
Qy        333 LDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVSCYVSD 392
              ||||:||:|:||:|||||||| |||||||||||||:| |||||||:|||||||||||||
Db        332 LDRLALRYDREGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVSCYVSD 391

Qy        393 DGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFVRERRA 452
              ||:|||||||||||:|||:||||||::|||||||:|| |||||||||:  :||:||||
Db        392 DGSAMLTFEALSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERRA 451

Qy        453 MKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSGGLDCE 512
              |||||||||||||||||||||||||||| ||| ||||| |||||||||||| ||||| :
Db        452 MKREYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGNNPRDHPGMIQVFLGHSGGLDTD 511

Qy        513 GNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAV 548
              |||||||||||||||||: ||||||||||:|||||
Db        512 GNELPRLVYVSREKRPGFQHHKKAGAMNALIRVSAV 547
```

RESULT 13
US-07-689-008-2
; Sequence 2, Application US/07689008
; Patent No. 5268274
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arie
; APPLICANT: Calhoon, Roger D
; APPLICANT: Fear, Anna L
; APPLICANT: Gelfand, David H
; APPLICANT: Meade, James H
; APPLICANT: Tal, Rony
; APPLICANT: Wong, Hing
; APPLICANT: Benziman, Moshe
; TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
; TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen
; STREET: Three Embarcadero Center
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,008
; FILING DATE: 19910422
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 337,194
; FILING DATE: 12-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 496,236
; FILING DATE: 23-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Lisabeth Feix
; REGISTRATION NUMBER: 31547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 393-2000
; TELEFAX: (415) 393-2286
; TELEX: 340817 MACPAG SFO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3031 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-689-008-2

Query Match 4.3%; Score 252; DB 1; Length 3031;
Best Local Similarity 20.4%; Pred. No. 8.2e-15;

Matches 168; Conservative 89; Mismatches 217; Indels 350; Gaps 36;

```
Qy        275 IIIIRLVVLGFFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQ---FPKWFP----- 326
              :|:  : || ||   |        | | ::| :  : : ::| | :   | |
Db         53 LIVALVCVLVFFIVGRGKSRRTQIF-LEVLSALVSLRY-LTWRLTETLDFDTWIQGGLGV 110

Qy        327 --IERETY-LDRLSLRFDKEGQPSQLAP----------IDFFVSTVDPLKEPPLVTTN 371
                :  | | |  | | : :  ||  ||           :| |: | |   |  :
Db        111 TLLMAELYALYMLFLSYFQTIQPLHRAPLPLPDNVDDWPTVDIFIPTYD---EQLSIVRL 167

Qy        372 TVLSILSVDYPVDKVSCYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYF 431
              |||  | :|:| |||: |: |||
Db        168 TVLGALGIDWPPDKVNVYILDDG---------------------------------- 190

Qy        432 QQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNN 491
                                :: |:|:|        ||
Db        191 ----------------------VRPEFEQF--------AK---------------------- 200

Qy        492 VRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTN 551
                                ||        :|: |      : | ||| :|  ::      |:
Db        201 ------------------DCGA------LYIGRVD----SSHAKAGNLNHAIK----RTS 228

Qy        552 APYLLNLDCDHYINNSKAIKEAMCFMMDPLLGKKVCYVQFPQRFDGIDRHDRYANRNVVF 611
              |:| ||||| |     :: || :|   :  :|:  :| |   |   |
Db        229 GDYILILDCDH-IPTRAFLQIAMGWM---VADRKIALMQTPHHFYSPD------------ 272

Qy        612 FDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNR 671
                                    |:|    ||     : ||                  ||
Db        273 ----------------------PFQRNLAVGY-----RTPPE----------------GN- 289

Qy        672 KQKKTTKPKTEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQS 731
                              ||:                              |::        |  :
Db        290 ---------------LFY--------------------------GVIQD---GNDFWDA 304

Qy        732 SVFVTSTLLENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWIYGSVTEDILTGF 791
              : |                  | | | :||| |                |:   :||||  |
Db        305 TFF------------CGSCAILRREAIESIG------------GFAVETVTEDAHTAL 338

Qy        792 KMHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGG 850
              :|   || : |  ||      | |   |:  : | :||| | |:||  ::  |    |||
Db        339 RMQRRGWSTAYLRIP---VASGLATERLTTHIGQRMRWARGMIQIFRVDNPML----GGG 391

Qy        851 LKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI---TPELNNVASLWFMSLFIC 907
              ||  :|  |::::   : :|| : :   |   |  |: |   :|                :
Db        392 LKLGQRLCYLSAMTSFFFAIPRVIFLASPLAFLFFGQNIIAASP-------------LA 437

Qy        908 IFATSILEMRWS---GVGIDDWWRNEQFW-VIGGVSSHLFAVFQGLLKVIAGVDTSFTVT 963
              : | :|  |  |      ::  ||  ||  :   :  || |   :: ::       | ||
Db        438 VLAYAIPHMFHSIATAAKVNKGWR-YSFWSEVYETTMALFLVRVTIITLMFPSKGKFNVT 496

Qy        964 SKGG--DDEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFG--K 1019
                |||  ::|||                        : |   | |  |  : | | |  :
Db        497 EKGGVLEEEEFD----------------------LGATYPNIIFAGIMTLGLLIGLFE 532

Qy       1020 LFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVWS-----ILLASI 1058
              | | |            | |   |   :  :|:     ||||:|
Db        533 LTFHF----------NQLAGIAKRAYLLNCIWAMISLIILLAAI 566
```

RESULT 14
US-08-960-048-11
; Sequence 11, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12

```
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 693
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-08-960-048-11

  Query Match             4.1%;  Score 241.5;  DB 3;  Length 693;
  Best Local Similarity   18.1%;  Pred. No. 7e-15;
  Matches  148;  Conservative   95;  Mismatches  205;  Indels  371;  Gaps   31;

Qy        274 MIIIIRLVVLGFFFHYRVM------HPVNDAFALWLISVICEIWFAMSWILDQFPKWFPI 327
              |:|:: | |   :  :|          ||:    | |:  |   |  :  :|  |   :|:
Db         21 MLIVLSLTVSCRYIWWRYTSTLNWDDPVSLVCGLILLFAITYAWIVL--VLGYFQVVWPL 78

Qy        328 ERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVS 387
              |:     : |  |    ||    :| || | :   |   | ||: : | :|:| ||::
Db         79 NRQP----VPLPKDMSLWPS----VDIFVPTYN---EDLNVVKNTIYASLGIDWPKDKLN 127

Qy        388 CYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFV 447
              :: |||
Db        128 IWILDDGG---------------------------------------------------- 135

Qy        448 RERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSG 507
                : |: :|          || |
Db        136 ------REEFRQF----------AQNV-------------------------------- 146

Qy        508 GLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNS 567
                  : |::|      : | ||| :|  :: :       ::   ||||    |
Db        147 ----------GVKYIARTT----HEHAKAGNINNALKYA----KGEFVSIFDCDHVPTRS 188

Qy        568 KAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID-------RHDRYANRNVVFFDINMKGLD 620
                |: :    |::  :| |  |   |        |  :  |   :|: :   | |
Db        189 FLQMTMGWFLKE----KQLAMMQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLVQDGND 244

Qy        621 GIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPK 680
                                          |    ||  |
Db        245 ------------------------------------MWDATFFCGSC------------ 255

Qy        681 TEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLL 740
                                                  :: :: |:
Db        256 -----------------------------------AVIRRKPLD-------------- 264

Qy        741 ENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMHCHGWR 799
                                             ||| |   :|||| |  ::|  |:
Db        265 ---------------------------------EIGGIAVETVTEDAHTSLRLHRRGYT 290

Qy        800 SIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFS 858
              | |  ||:     | |  :||  : | :||| | ::||  :: ||    | |||| :|
Db        291 SAYMRIPQAA---GLATESLSAHIGQRIRWARGMVQIFRLDN-PL---TGKGLKFAQRLC 343

Qy        859 YINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFA------TS 912
              |:|:: :  : || | : | |   ||   :|         ::  :| | :|       |
Db        344 YVNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---------IYAPALMIALFVLPHMIHAS 394

Qy        913 ILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGG 967
              :   : |     :| :| :     | |   :         |: :|       | ||:|||
Db        395 LTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT---------LVALINPHKGKFNVTAKGG 444

Qy        968 D--DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFW 1025
              :||: :     |  :  |   |:||| :||  |:          |   :|
Db        445 GLVEEEYVD-----W-VISRPYIFLVLLNLVGVAVGI----------WRYFYG------- 481

Qy       1026 VIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
                              ||       :|  : |::||
Db        482 -----------------PPT------EMLTVVVSMVWV 496


RESULT 15
US-09-838-586-11
; Sequence 11, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
;  APPLICANT: Stalker, D. et al.
;  TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
;  TITLE OF INVENTION:  Sequences
;  FILE REFERENCE: 15621/02/US
;  CURRENT APPLICATION NUMBER: US/09/838,586
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: 60/029,987
;  PRIOR FILING DATE: 1996-10-29
;  PRIOR APPLICATION NUMBER: 08/960,048
;  PRIOR FILING DATE: 1997-10-29
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 693
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-838-586-11

  Query Match             4.1%;  Score 241.5;  DB 4;  Length 693;
  Best Local Similarity   18.1%;  Pred. No. 7e-15;
  Matches  148;  Conservative   95;  Mismatches  205;  Indels  371;  Gaps   31;

Qy        274 MIIIIRLVVLGFFFHYRVM------HPVNDAFALWLISVICEIWFAMSWILDQFPKWFPI 327
              |:|:: | |   :  :|          ||:    | |:  |   |  :  :|  |   :|:
Db         21 MLIVLSLTVSCRYIWWRYTSTLNWDDPVSLVCGLILLFAITYAWIVL--VLGYFQVVWPL 78

Qy        328 ERETYLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSILSVDYPVDKVS 387
              |:     : |  |    ||    :| || | :   |   | ||: : | :|:| ||::
Db         79 NRQP----VPLPKDMSLWPS----VDIFVPTYN---EDLNVVKNTIYASLGIDWPKDKLN 127

Qy        388 CYVSDDGAAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPEWYFQQKIDYLKDKVAANFV 447
              :: |||
Db        128 IWILDDGG---------------------------------------------------- 135

Qy        448 RERRAMKREYEEFKVRINALVAKAQKVPEEGWTMQDGTPWPGNNVRDHPGMIQVFLGQSG 507
                : |: :|          || |
Db        136 ------REEFRQF----------AQNV-------------------------------- 146

Qy        508 GLDCEGNELPRLVYVSREKRPGYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNS 567
                  : |::|      : | ||| :|  :: :       ::   ||||    |
Db        147 ----------GVKYIARTT----HEHAKAGNINNALKYA----KGEFVSIFDCDHVPTRS 188

Qy        568 KAIKEAMCFMMDPLLGKKVCYVQFPQRFDGID-------RHDRYANRNVVFFDINMKGLD 620
                |: :    |::  :| |  |   |        |  :  |   :|: :   | |
Db        189 FLQMTMGWFLKE----KQLAMMQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLVQDGND 244

Qy        621 GIQGPIYVGTGCVFRRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKQKKTTKPK 680
                                          |    ||  |
Db        245 ------------------------------------MWDATFFCGSC------------ 255

Qy        681 TEKKKLLFFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQQKLEKKFGQSSVFVTSTLL 740
                                                  :: :: |:
Db        256 -----------------------------------AVIRRKPLD-------------- 264

Qy        741 ENGGTLKSASPASLLKEAIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMHCHGWR 799
                                             ||| |   :|||| |  ::|  |:
Db        265 ---------------------------------EIGGIAVETVTEDAHTSLRLHRRGYT 290

Qy        800 SIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLWYGYGGGLKFLERFS 858
              | |  ||:     | |  :||  : | :||| | ::||  :: ||    | |||| :|
Db        291 SAYMRIPQAA---GLATESLSAHIGQRIRWARGMVQIFRLDN-PL---TGKGLKFAQRLC 343

Qy        859 YINSIVYPWTSIPLLAYCTLPAICLLTGKFITPELNNVASLWFMSLFICIFA------TS 912
              |:|:: :  : || | : | |   ||   :|         ::  :| | :|       |
Db        344 YVNAMFHFLSGIPRLIFLTAPLAFLLLHAYI---------IYAPALMIALFVLPHMIHAS 394

Qy        913 ILEMRWSGVGIDDWWRNEQF-----WVIGGVSSHLFAVFQGLLKVIAGVDTSFTVTSKGG 967
              :   : |     :| :| :     | |   :         |: :|       | ||:|||
Db        395 LTNSKIQGKYRHSFW-SEIYETVLAWYIAPPT---------LVALINPHKGKFNVTAKGG 444

Qy        968 D--DEEFSELYTFKWTTLLIPPTTLLLLNFIGVVAGVSNAINNGYESWGPLFGKLFFAFW 1025
              :||: :     |  :  |   |:||| :||  |:          |   :|
Db        445 GLVEEEYVD-----W-VISRPYIFLVLLNLVGVAVGI----------WRYFYG------- 481
```

```
Qy       1026 VIVHLYPFLKGLVGRQNRTPTIVIVWSILLASIFSLLWV 1064
                                ||        :|  : |::||
Db        482 ------------------PPT------EMLTVVVSMVWV 496
```

Search completed: December 15, 2003, 13:05:32
Job time : 27 secs